

FERNANDA PACHECO

**SEQUENCIAMENTO, IDENTIFICAÇÃO E ANÁLISE DE GENES DE
ARROZ ENVOLVIDOS NA INTERAÇÃO COM *Herbaspirillum*
seropedicae.**

Dissertação apresentada como
requisito parcial para obtenção do
título de Mestre em Bioquímica,
Setor de Ciências Biológicas,
Universidade Federal do Paraná.

Orientador: Emanuel Maltempi de
Souza

Coorientadores: Liu Un Rigo
Luciano Fernandes Huergo

Curitiba

2008

AGRADECIMENTOS

À uma energia superior que me deu forças durante esta caminhada, permitindo que eu nunca desistisse.

Ao professor Fábio Oliveira Pedrosa pela oportunidade de trabalhar neste grupo.

Ao meu orientador Emanuel Maltempi de Souza pelo apoio, confiança, correções e paciência, sem o qual não teria sido possível meu crescimento profissional.

Aos coorientadores Liu Un Rigo e Luciano Fernandes Huergo pelo auxílio com as correções, paciência e amizade.

Ao Salah, pela amizade, apoio, carinho, além de ser o mentor deste trabalho.

À Coordenadoria de Aperfeiçoamento Pessoal de Ensino Superior (CAPES) e ao Conselho nacional de Pesquisa e Desenvolvimento (CNPq) pelo apoio financeiro.

Aos professores do Núcleo de Fixação Biológica de Nitrogênio do Departamento de Bioquímica que direta ou indiretamente contribuíram para a produção deste trabalho.

À professora Rose Adele pela leitura e correções deste trabalho, de toda a sua boa vontade com o andamento das coisas no laboratório, além de seu carinho pelos animais.

Ao professor Leonardo Cruz pelo suporte técnico, dedicação e amizade.

À Lílian Noindorf e Gisele Klassen pela leitura e correções deste trabalho, além da paciência e compreensão.

Aos funcionários Valter Baura, Julieta Pie e Roseli Prado pelo apoio técnico, acolhimento, boa vontade, cooperação e amizade.

Aos colegas de mestrado e aos demais colegas de laboratório pelo apoio, companheirismo, momentos difíceis e também pelos momentos de alegria.

Aos novos amigos Doumit, Danilo, Rafael, Dani, Carol, Daniela Seixas, Lilian, Juliana Inaba, Giovana, Arnaldo, Patilene, Ana Cláudia, Larissa C. e Liziane que amenizaram os momentos difíceis e que tornaram momentos comuns da vida em momentos felizes, ou pelo menos engraçados.

Ao Héllisson, Arnaldo, Marco K., Juliana Osaki, Leonardo, Giovani e Liziane, que além da amizade, foram imprescindíveis no suporte técnico, cedendo além de seu tempo, amizade e compreensão.

À Tâmara, Angélica, Clau e Jana, que estiveram sempre presentes de alguma forma ao longo de vários anos na minha vida.

Ao Jonas, pelo carinho, dedicação, amor e compreensão.

À minha família que sempre me apoiou e confiou em mim, pelo carinho e dedicação e por terem fornecido alicerce à minha vida e ao meu caráter.

SUMÁRIO

1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	4
2.1 FIXAÇÃO BIOLÓGICA DE NITROGÊNIO	4
2.2 NITROGENASE	4
2.3 <i>Herbaspirillum seropedicae</i>	5
2.4. INTERAÇÃO PLANTA-BACTÉRIA	6
2.4.1 Tipos de interação	6
2.4.2 Vantagens de permanência no interior da planta e colonização	7
2.4.3 Benefícios da Colonização	8
2.4.5 Sinalização molecular planta-bactéria	10
2.4.6 Associação endófitos-gramíneas	11
2.5. ESTs	12
2.6 SEQUENCIAMENTO DO GENOMA DE ARROZ	14
3. OBJETIVOS	15
3.1 OBJETIVOS GERAIS	15
3.2 OBJETIVOS ESPECÍFICOS	15
4. MATERIAL E MÉTODOS	16
4.1 BACTÉRIAS E VETOR PLASMIDIAL DE EXPRESSÃO	16
4.2 OBTENÇÃO DE ESTS (BRUSAMARELO, 2007)	16
4.3 MANIPULAÇÃO DE DNA	19
4.3.1 Extração de DNA plasmidial em placas de 96 poços	19
4.3.2 Eletroforese de DNA	20
4.3.3 Reação de Sequenciamento	20
4.4 ANÁLISE DE BIOINFORMÁTICA DAS ESTs	21
4.4.1 Nomenclatura utilizada para nomear as seqüências de ESTs	22
4.4.2 Edição, montagem e alinhamento das seqüências	23
4.4.3 Busca por homologia em bancos de dados	25
5. RESULTADOS E DISCUSSÃO	27
5.1 SEQUENCIAMENTO	28
5.2 BUSCA POR SIMILARIDADE EM BANCOS DE DADOS	30
5.2.1 Categorização funcional	33
5.3 IDENTIFICAÇÃO DOS GENES PROVAVELMENTE EXPRESSOS DIFERENCIALMENTE	35
5.3.1 Genes encontrados exclusivamente na biblioteca inoculada com <i>H.seropedicae</i>	36
5.3.2 Genes encontrados preferencialmente na biblioteca inoculada com <i>H.seropediace</i>	39
5.3.3 Genes mais freqüentes encontrados em ambas bibliotecas	47
5.3.4 Prováveis novos genes encontrados neste estudo	52
5.4 CONSIDERAÇÕES FINAIS	54
6. CONCLUSÕES	55
REFERÊNCIAS BIBLIOGRÁFICAS	56

LISTA DE FIGURAS

FIGURA 1 - FLUXOGRAMA DE OBTENÇÃO DE ESTS DE <i>Oryza sativa</i>	18
FIGURA 2 – FLUXOGRAMA GERAL DA ANÁLISE DAS ESTS IN SILICO:.....	21
FIGURA 3 – FLUXOGRAMA DETALHADO DA ANÁLISE DAS ESTS	22
FIGURA 4 – NOMENCLATURA UTILIZADA PARA OS.....	23
ELETROFORETOGRAMAS.....	23

LISTA DE GRÁFICOS

GRÁFICO 1 – DISTRIBUIÇÃO DAS SEQUÊNCIAS DE ESTs DA BIBLIOTECA RRCH5SN DE ACORDO COM SEU TAMANHO E QUALIDADE.....	29
GRÁFICO 2 – DISTRIBUIÇÃO DAS SEQUÊNCIAS DE ESTs DA BIBLIOTECA RRSN DE ACORDO COM SEU TAMANHO E QUALIDADE.....	30
GRÁFICO 3 - DISTRIBUIÇÃO DE SEQUÊNCIAS DA BIBLIOTECA RRSN DE cDNA DE RAÍZES DE ARROZ NOS REFERENTES CROMOSSOMOS DE <i>Oryza sativa</i>	32
GRÁFICO 4 - DISTRIBUIÇÃO DE SEQUÊNCIAS DA BIBLIOTECA RRCH5SN DE cDNA DE RAÍZES DE ARROZ NOS REFERENTES CROMOSSOMOS DE <i>Oryza sativa</i>	33
GRÁFICO 5 – CLASSIFICAÇÃO FUNCIONAL DA BIBLIOTECA RRSN BASEADA NO <i>GENE ONTOLOGY</i>	34
GRÁFICO 6 – CLASSIFICAÇÃO FUNCIONAL DA BIBLIOTECA RRCH5SN BASEADA NO <i>GENE ONTOLOGY</i>	35

LISTA DE TABELAS

TABELA 1 – BACTÉRIAS E VETOR UTILIZADO:.....	16
TABELA 2 – SCRIPTS UTILIZADOS PARA PROCESSAMENTO DOS ELETROFORETOGRAMAS	24
TABELA 3 – PROGRAMAS DE BIOINFORMÁTICA UTILIZADOS NA MONTAGEM E ANÁLISE DAS SEQUÊNCIAS DE EST	25
TABELA 4 - CARACTERÍSTICAS DAS BIBLIOTECAS DE cDNA DE RAÍZES DE ARROZ UTILIZADAS	27

LISTA DE ABREVIATURAS

BASH - linguagem para execução de programas em ambiente Linux (*Born Again Shell*)

BLAST - Ferramenta de alinhamento local, do inglês *Basic Local Alignment Search Tool*

CAP3 - Contig Assembly Program

Cm - Cloranfenicol

cDNA - DNA complementar

DNA - ácido desoxirribonucléico

EST - Etiqueta de sequencias expressas, do inglês *Expressed Sequence Tags*

mRNA - RNA mensageiro

nr - sequência de proteínas não redundante ,do inglês *non-redundant protein sequences*

RNA - ácido ribonucléico

TB - Terrific Broth

UTR - Regiões não traduzidas, do inglês *Untranslated Regions*

RESUMO

Arroz (*Oryza sativa*) é o principal alimento para mais da metade da população mundial. O nitrogênio é um dos fatores que mais limitam sua produção que, devido à crescente demanda mundial, requer utilização de grandes quantidades de fertilizantes industriais nitrogenados. Estes insumos, além de serem economicamente onerosos, também causam danos ao meio ambiente. Uma alternativa para este uso seria a exploração da fixação biológica de nitrogênio por bactérias diazotróficas endofíticas como, por exemplo, *Herbaspirillum seropedicae*. Para isto, é necessário compreender melhor os mecanismos moleculares envolvidos na interação planta-bactéria. O sequenciamento de ESTs de gramíneas inoculadas com bactérias endofíticas fixadoras de nitrogênio tem se revelado uma importante ferramenta para a compreensão da interação planta-bactéria. Com este objetivo foram seqüenciadas biblioteca de cDNA de raízes de plântulas de arroz obtidas 5 dias após inoculação com *H. seropedicae* e de raízes de plântulas não inoculadas. Um total de 4608 clones tiveram suas extremidades 5' seqüenciadas. Foram 1377 seqüências de ESTs únicas com qualidade Phred igual ou superior a 15. Foram identificados genes expressos exclusivamente na biblioteca inoculada com *H. seropedicae*, dentre os quais: transportadores de nitrato, genes codificando proteínas contendo o domínio WD40, com envolvimento na regulação da biossíntese de flavonóides, e um gene codificando uma proteína SEP2 de resposta a estresse. Genes de resposta a fitormônios como etileno e auxina foram encontrados expressos preferencialmente na biblioteca de cDNA de arroz inoculada com *H. seropedicae*. Também foram encontrados 5 prováveis novos genes de arroz. A investigação da função biológica destas ESTs pode conduzir à identificação de novos genes envolvidos nesta associação e à elucidação dos sinais moleculares envolvidos na interação arroz-*H. seropedicae*.

Palavras-chave: *Oryza sativa*, ESTs, *Herbaspirillum seropedicae*, cDNA, interação.

1. INTRODUÇÃO

O arroz (*Oryza sativa*) é uma gramínea da família Poaceae, da classe Liliopsida, da ordem Poales, pertencente ao grupo das monocotiledôneas, contido na divisão das Magnoliófitas. Segundo WATSON e DALLWITZ (1992) a família Poaceae contém 650 gêneros e aproximadamente 9000 espécies, incluindo vários cereais de importância econômica como milho, trigo, sorgo, cevada, além da cana-de-açúcar.

Os tamanhos dos genomas das Poaceae variam consideravelmente (GOFF et al., 2002), mas apresentam alta sintenia e as seqüências dos genes possuem alta similaridade. Os genomas de sorgo, milho, centeio e trigo possuem tamanho estimado em 1000, 3000, 5000 e 16000 Mb, respectivamente (GOFF et al., 2002). O arroz tem um dos menores genomas, aproximadamente 389 Mb (GOFF et al., 2002). O pequeno genoma e a alta densidade de prováveis genes fazem do arroz um alvo atrativo para descoberta de genes em cereais e análise genômica (KIKUCHI et al., 2003; GOFF et al., 2002).

Arroz, trigo e milho são os três cereais mais importantes. Mais de 500 milhões de toneladas de cada um são produzidas anualmente (WASDE, 2001). A maioria do arroz produzido é diretamente consumida por humanos, provendo alimento básico para mais da metade da população mundial (IRRI, 2006).

A produção de cereais vem acompanhando o crescimento populacional (BOCKMAN et al. 1990). Desnutrição e fome estão ainda difundidas em muitas regiões do mundo devido principalmente a fatores políticos, como pobreza e guerras, e não pela produção agrícola ter atingido um limite biológico máximo para produção de alimentos (BOCKMAN et al. 1997). Entretanto como a população mundial permanece em crescimento, provavelmente alcançando aproximadamente 8,3 bilhões em 2025 antes de atingir um nível estável, existe uma necessidade premente de aumento da produção de alimentos e provável expansão de áreas para agricultura intensiva (BOCKMAN et al. 1997). Calcula-se que para continuar suprimindo a demanda alimentar das próximas três décadas, será necessário

aumentar a produção em 60% em relação à atual (LADHA e REDDY, 2003). O aumento na produção de alimentos requer a interação de vários fatores, dentre eles, o aumento na disponibilidade de nutriente à planta. Com exceção da água, o nitrogênio é o nutriente que mais freqüentemente limita a produção do arroz (LADHA e REDDY, 2003; BOCKMAN et al. 1997; STOLTZFUS et al., 1997). Sendo assim, a agricultura moderna utiliza intensivamente fertilizantes industriais nitrogenados para assegurar alta produtividade (CHOUDHURY e KENNEDY, 2004; BOCKMAN et al. 1997). O crescente uso de fertilizantes químicos constitui uma grande interferência humana no ciclo do nitrogênio (DIXON e KAHN, 2004).

A produção e uso de fertilizantes nitrogenados têm um elevado consumo de combustíveis fósseis, que além de ser um recurso não renovável, apresenta alto custo e riscos ao ambiente (LADHA e REDDY, 2003; BOCKMAN et al. 1997; STOLTZFUS et al., 1997). O processo de produção destes compostos nitrogenados produz CO_2 que contribui para o aquecimento global devido ao seu efeito estufa. Quando aplicados no solo, os efeitos ao ambiente são agravados, pois mais da metade do nitrogênio aplicado é perdido através de desnitrificação, volatilização da amônia e escoamento para o solo (BOCKMAN et al. 1997; STOLTZFUS et al., 1997). A perda do nitrogênio não assimilado implica em contaminação dos lençóis freáticos por nitrato, representando riscos à saúde (LADHA e REDDY, 2003; STOLTZFUS et al., 1997), bem como a acidificação do solo e eutrofização da água (DIXON e KAHN, 2004; KENNEDY e TCHAN, 1992). Além disso, parte do excesso de fertilizante nitrogenado aplicado é transformada em óxido nitroso (N_2O), um gás de potente efeito estufa e que também contribui para a destruição da camada de ozônio (LADHA e REDDY, 2003; STOLTZFUS et al, 1997). Adicionalmente a estes problemas ambientais, o uso de uréia em longo prazo exaure a matéria orgânica do solo (CHOUDHURY e KENNEDY, 2004).

A fixação biológica de nitrogênio é uma alternativa para produção agrícola sustentável. A utilização de bactérias diazotróficas endofíticas promotoras do crescimento vegetal poderia ser explorada (LADHA e REDDY, 2003; STOLTZFUS

et al, 1997; WU et al., 1995) com ganhos ecológicos e econômicos (MUTHUKUMARASAMY et al., 1995).

Herbaspirillum seropedicae é um fixador de nitrogênio que coloniza gramíneas e fixa nitrogênio de forma mais eficiente que outros diazotrofos como *Gluconacetobacter diazotrophicus* e *Burkholderia* spp. quando inoculados em cana-de-açúcar e arroz, respectivamente (CANUTO et al., 2003; BALDANI, BALDANI e DÖBEREINER, 2000).

Este trabalho tem como principal objetivo compreender os mecanismos de interação planta-bactéria (*Oryza sativa*-*Herbaspirillum seropedicae*) por meio de análise de genes de arroz expressos quando este é colonizado por *H. seropedicae*. Espera-se que a identificação destes genes permita compreender o mecanismo molecular de interação *H. seropedicae* com arroz e outras gramíneas de interesse econômico.

2. REVISÃO DE LITERATURA

2.1 FIXAÇÃO BIOLÓGICA DE NITROGÊNIO

O nitrogênio é o elemento necessário para síntese de biomoléculas como proteínas, ácidos nucleicos, dentre outras, sendo, portanto, fundamental para a manutenção de todas as formas de vida no planeta. O gás dinitrogênio (N₂) é a forma mais abundante deste elemento e constitui aproximadamente 80% da atmosfera da Terra, porém nenhum eucarioto é capaz de utilizá-lo diretamente (POSTGATE,1982).

Apenas algumas espécies de procariontes são capazes de utilizar N₂ através de um processo denominado fixação biológica de nitrogênio. Neste processo, o nitrogênio atmosférico é convertido em amônio, uma forma metabolicamente assimilável pelas plantas. Esta reação é catalisada pelo complexo enzimático da nitrogenase (BURRIS,1991). As bactérias capazes de realizar esta conversão são denominadas diazotrofos (BURRIS,1991).

A fixação de nitrogênio é uma importante parte do ciclo do nitrogênio e funciona como repositório do nitrogênio assimilável na biosfera, compensando as perdas provenientes da desnitrificação (DIXON e KAHN, 2004).

2.2 NITROGENASE

A nitrogenase é a enzima responsável pela fixação biológica de nitrogênio, catalisando a redução do dinitrogênio gasoso a amônio, conforme a reação (SIMPSON e BURRIS 1984):



O complexo da nitrogenase contém dois componentes. O componente menor é o homodímero γ_2 , conhecido como proteína ferro, dinitrogenase redutase

ou NifH, que funciona como um doador de elétrons (dependente de ATP) para o componente heterodimérico maior, a proteína ferro-molibdênio, dinitrogenase ou NifDK, que contém o sítio catalítico da enzima (HOWARD e REES, 1996; DIXON e KAHN, 2004). Ambas as proteínas componentes da nitrogenase são extremamente sensíveis ao oxigênio (DIXON e KAHN, 2004).

Além dos genes estruturais *nifHDK*, que codificam as proteínas NifH e NifDK, a biossíntese e atividade da nitrogenase dependem de produtos codificados por outros genes *nif*, essenciais para o transporte de elétrons, regulação da transcrição e maturação dos componentes da nitrogenase (MERRICK, 1992). Em vista do alto gasto energético envolvido na redução de N_2 a NH_4^+ , o processo de fixação de nitrogênio é altamente regulado tanto ao nível de atividade da nitrogenase como na expressão gênica (POSTGATE, 1982).

2.3 *Herbaspirillum seropedicae*

A capacidade de fixar nitrogênio é encontrada na maioria dos grupos filogenéticos bacterianos. Estes organismos denominados diazotrofos, são capazes de crescer utilizando nitrogênio atmosférico como única fonte de nitrogênio, e são encontrados em uma ampla variedade de habitats, tanto livres no solo ou na água ou associados em simbiose com plantas (DIXON e KAHN, 2004).

H. seropedicae é uma bactéria diazotrófica, endofítica, gram-negativa, vibrióide, membro da subdivisão β das Proteobactérias, capaz de fixar nitrogênio em condições microaeróbicas e em uma ampla faixa de pH (5,3 a 8,0) (BALDANI et al., 1986). Esta bactéria foi encontrada em raízes, caules e folhas de espécies de gramíneas, incluindo várias economicamente importantes como milho, arroz, trigo, sorgo e cana-de-açúcar (RONCATO-MACARI et al., 2003a; BALDANI et al., 1986), e ainda em espécies tropicais como banana e abacaxi (CRUZ et al, 2001).

H. seropedicae não sobrevive bem no solo (OLIVARES, 1996; BALDANI, 1986) e não causa sintomas de doenças quando associado às plantas (OLIVARES, 1996; PIMENTEL, 1991; BALDANI, 1986). Fixa nitrogênio sob condições

microaeróbicas (com baixa concentração de O_2) e cresce com N_2 como única fonte de nitrogênio (BALDANI et al., 1986).

2.4.INTERAÇÃO PLANTA-BACTÉRIA

2.4.1 Tipos de interação

Interações planta-bactéria podem ser subdivididas em três classes: simbiótica, patogênica e associativa (PÜHLER et al, 2004). Independente do tipo de interação planta-bactéria, a sobrevivência e a persistência do microrganismo em diversas partes do hospedeiro vegetal necessitam de respostas adaptativas por parte da bactéria (MILLER et al, 1989).

Na prática, todas as plantas vivem em íntima associação com microrganismos, que podem colonizar a superfície das plantas, caracterizando colonização epifítica; ou ocupando espaços dentro dos tecidos das plantas, sendo denominada colonização endofítica (BRENCIC e WINANS, 2005).

H. seropedicae, microrganismo de estudo deste trabalho, é um microrganismo de associação endofítica, que se caracteriza por passar a maior parte da sua vida no interior das plantas (QUISPEL, 1992). Este organismo possui ainda a característica de fixar nitrogênio atmosférico é então denominado diazotrofo endofítico. *H. seropedicae* foi originalmente isolado de raízes de cereais (BALDANI et al, 1986), e também foi encontrado em caules e folhas de arroz e cana-de-açúcar (DOBEREINER, 1992).

Diazotrofos endofíticos como o *H. seropedicae* apresentam estreita associação com o hospedeiro e baixa sobrevivência em solos desprovidos de raízes (BALDANI, 1996). Ao colonizarem os espaços intercelulares e vasos das plantas associadas, estes organismos são capazes de se dispersar sistemicamente até alcançar tecidos aéreos (JAMES e OLIVARES, 1998; BALDANI et al., 1986).

2.4.2 Vantagens da colonização interna

O interior da planta oferece às bactérias endofíticas a vantagem adicional de proteção e redução de competição por nutrientes (STOLTZFUS, 1997). Portanto, os organismos que colonizam o interior da planta hospedeira estão protegidos das mudanças ambientais que sofrem as bactérias em contato com o solo ou presentes na rizosfera (REINHOLD-HUREK e HUREK, 1998). Por habitar o interior das plantas, estas bactérias evitam competição por substratos de carbono com bactérias da rizosfera obtendo nutrientes diretamente da planta hospedeira (LADHA e REDDY; 2003; JAMES e OLIVARES, 1998; BODDEY et al, 1995, DÖBEREINER et al, 1995). Adicionalmente, o interior da planta, com baixa concentração de O₂ e disponibilidade de fontes de carbono, provê um ambiente adequado para fixação de nitrogênio (JAMES e OLIVARES, 1998; SPRENT e JAMES, 1995). Células de *H.seropedicae* expressando genes *nif* foram encontradas no tecido vascular nos caules de sorgo, trigo, milho e arroz (RONCATO-MACCARI et al., 2003a).

A penetração de *H. seropedicae* em plantas de arroz aparentemente envolve a adesão às raízes primeiramente na zona de alongação e diferenciação, seguida de colonização dos pontos de emergência de raízes secundárias e penetração através da epiderme principalmente em pontos de emergência da raiz lateral. A bactéria então coloniza e ocupa os espaços intercelulares, aerênquima e xilema e partes aéreas (RONCATO-MACCARI et al., 2003a; 2003b; JAMES et al., 2002; GYANESHWAR et al, 2002). Um número relativamente menor de bactéria foi encontrado nas partes aéreas quando comparado ao número associado às raízes (RONCATO-MACCARI et al., 2003).

O uso de inoculantes de diazotrofos endofíticos pode ser uma das formas de aproveitar o potencial da interação planta-microrganismo. De fato, acredita-se que diazotrofos endofíticos, particularmente *Gluconacetobacter diazotrophicus* e *Herbaspirillum* spp, possam ser responsáveis pela transferência de uma quantidade significativa de nitrogênio fixado em algumas variedades de cana de açúcar cultivadas no Brasil (*Saccharum* spp.) (JAMES e OLIVARES, 1998;

URQUIAGA, et al., 1992). Além disto, é possível que variedades de arroz tenham potencial para aproveitar a fixação de nitrogênio e possam se beneficiar da interação com endófitos (SHRESTHA et al., 1996). Sugere-se que a associação entre plantas de arroz e bactérias fixadoras de nitrogênio seja controlada pelo genótipo da planta de arroz (CHOUDHURY e KENNEDY, 2004). Variedades tolerantes a alumínio apresentaram atividade de redução de acetileno significativamente maior do que uma semi-tolerante quando inoculada com *H. seropedicae* Z67 (GYANESHWAR et al, 2002). As diferenças nas atividades observadas talvez não sejam devido a diferenças no número de bactérias associadas às plantas, pois um número similar foi isolado de todas as variedades, sugerindo que as bactérias associadas com variedades alumínio-tolerante foram fisiologicamente mais ativas que aquelas associadas com variedades sensíveis, levando a um aumento do crescimento e acúmulo de nitrogênio. Bactérias associadas com as variedades tolerantes expressaram maiores níveis da proteína Fe (NifH) e a atividade de redução de acetileno encontrava-se aumentada em relação a outras variedades, caracterizando aumento na atividade da nitrogenase (GYANESHWAR et al, 2002)

O efeito das bactérias fixadoras de nitrogênio sobre a planta associada depende também da estirpe inoculada. BALDANI, BALDANI e DÖBEREINER (2000) verificaram que o aumento no conteúdo de nitrogênio da planta induzido pela fixação biológica de nitrogênio variou de 31-54% dependendo da estirpe utilizada.

2.4.3 Benefícios da Colonização

A associação diazotrofo-gramínea geralmente é benéfica a ambos os parceiros. Diazotrofos geralmente não causam danos ao organismo hospedeiro, e ainda promovendo crescimento da planta pela produção e secreção de reguladores de crescimento, atividade antagonista a fitopatógenos e provendo nitrogênio biologicamente fixado (BENHAMOU et al, 1996; VAN BUREN et al., 1993). Auxina,

ácido indol acético e giberelinas são exemplos de hormônios reguladores de crescimento produzidos pelos endófitos diazotrofos (BÁSTIAN et al. 1998; FUENTES-RAMIREZ et al., 1993). Contudo, a intensidade do estímulo de crescimento vegetal, incluindo a fixação de nitrogênio *in planta* e a transferência deste nitrogênio fixado para a planta, depende de uma interação planta-bactéria eficiente (BULL et al., 1991; VAN PEER e SCHIPPER, 1989).

Plantas de arroz podem estabelecer associações com várias bactérias fixadoras de nitrogênio, pertencentes a vários grupos (BARRAQUIO, et al., 1982; MALIK et al., 1997). Algumas destas bactérias podem ser responsáveis pelo suprimento de nitrogênio fixado à planta (LADHA e REDDY, 1995). Além dos endófitos cultiváveis associados ao arroz, uma grande diversidade de fixadores de nitrogênio foi encontrada através de análise de seqüências parciais dos genes *nifH* ou *nifD* obtidas de amplificado de DNA de raízes (UEDA et al., 1995).

Um exemplo bem estudado de fixação de N₂ é a interação simbiótica rizóbio-legume, em que a bactéria fixa nitrogênio como endossimbionte. Esta simbiose é baseada em reconhecimento específico de sinais moleculares, que são produzidos tanto pela planta como pela bactéria (SPAINK, 2000). Sob condições ambientais apropriadas, rizóbios e plantas hospedeiras podem iniciar uma interação simbiótica, resultando no desenvolvimento de nódulos nas raízes. O interior do nódulo constitui um microambiente rico em nutrientes cuja concentração de oxigênio é controlada, e a bactéria habita os nódulos como endossimbionte fixador de nitrogênio. Este processo envolve uma troca altamente coordenada de sinais entre a planta e a bactéria e conduz a uma diferenciação gradual e coordenada e ajuste da fisiologia e metabolismo de ambos os parceiros da interação (BLADERGROEN e SPAINK. 1998; BROUGHTON et al, 2000; PERRET, 2000; SCHULTZE e KONDOROSI. 1998; SPAINK, H. P. 2000).

As gramíneas não formam estruturas simbióticas especializadas, como nódulos nas raízes (REINHOLD-HUREK e HUREK, 1998). Entretanto, bactérias diazotróficas têm sido encontradas dentro de tecidos de algumas gramíneas, e tem sido mostrado que algumas culturas de gramíneas, como variedades de cana-de-

açúcar cultivadas no Brasil sem adubação nitrogenada podem derivar uma parcela substancial do seu requerimento de nitrogênio da fixação de N₂ (BODDEY et al., 1995). Observações como estas continuam incentivando estudos para determinar como a bactéria habitando tecidos internos de gramíneas pode contribuir diretamente com o nitrogênio fixado para seu hospedeiro (BALDANI, 1996).

2.4.5 Sinalização molecular planta-bactéria

Durante o desenvolvimento das associações com bactérias, o sistema de defesa da planta é estritamente regulado, determinando se a interação será bem sucedida (LAMBAIS et al, 2001). Tem sido proposto que durante as interações planta-microrganismo o sistema de defesa da planta seja regulado por sinalização complexa via comunicação cruzada e transdução de sinais (MC DOWELL e DANGL, 2000; EUGEM et al, 1999). Por exemplo, nas interações planta-patógeno a ativação do sistema de defesa é dependente do reconhecimento de determinantes de virulência por receptores das plantas. Nestas condições, sinais específicos das vias de sinalização são disparados resultando em reações de defesa da planta (LAMBAIS et al, 2001; MC DOWELL e DANGL, 2000).

Assim, sejam as interações entre microrganismo e hospedeiro eucariótico, patogênicas ou não, dependem de uma perfeita comunicação molecular entre os parceiros. A detecção destes sinais conduz a padrões alterados de expressão de genes que culmina em mudanças específicas e adaptativas na fisiologia de ambos e que são requeridas para estas associações (BRENCIC e WINANS, 2005).

Durante o curso da interação planta-bactéria, a bactéria continua a monitorar as mudanças na fisiologia de seu hospedeiro. Estas mudanças são freqüentemente devidas a atividades específicas dos microrganismos colonizadores, que em resposta fazem ajustes contínuos na sua própria fisiologia. Desse modo, detecção e resposta de vários sinais do hospedeiro na interação planta-microrganismo é um

processo contínuo. Em muitos casos, adicionalmente a proteínas regulatórias específicas, reguladores globais apresentam função nestas interações (BRENCIC e WINANS, 2005).

Talvez uma das mais importantes adaptações em bactérias gram-negativas capazes de se associar a hospedeiros eucarióticos seja o sistema de secreção do tipo III (T3SS), uma estrutura bacteriana que evoluiu especificamente para entregar proteínas bacterianas nas células eucarióticas. Embora identificado originalmente em um grande número de bactérias patogênicas, o T3SS ocorre em um grande número de espécies de bactérias que são simbióticas ou ocorrem em associações não patogênicas com humanos, outros animais incluindo insetos ou nematóides, ou plantas (GALÁN e WOLF-WATZ, 2006).

2.4.6 Associação endófitos-gramíneas

Embora o aparente benefício e as características não-patogênicas da associação entre gramínea e *Herbaspirillum*, é razoável esperar que a planta possua mecanismos de reconhecimento da bactéria e de modulação da resposta de defesa contra o endófito, bem como exerça um controle da colonização interna pela bactéria. Sendo assim, a planta deve ter uma participação ativa na interação, respondendo a diversos processos metabólicos durante a associação (NOGUEIRA et al, 2001).

O controle da colonização da planta por bactérias endofíticas provavelmente envolve a interação molecular de produtos gênicos da planta e da bactéria. Cana-de-açúcar por exemplo, apresenta diferentes níveis de expressão de genes como POX9 (peroxidase), PAL2 (fenilamina liase) e CAT3 (catalase) quando inoculada com diferentes endófitos (*Gluconacetobacter diazotrophicus* e *H. rubrisubalbicans*) (LAMBAIS et al, 2001). Sugere-se que o processo de controle da infecção do hospedeiro e/ou colonização dependem tanto da estirpe do endófito (LAMBAIS et al, 2001), quanto a atividade de fixação de nitrogênio depende da variedade do

arroz com a qual o diazotrofo interage (GYANESHWAR et al, 2002; SANO et al., 1981).

Apesar do efeito benéfico já demonstrado, a natureza da colonização dos tecidos pelos diazotrofos endofíticos não está claramente compreendida, principalmente devido à dispersão da bactéria por todas as estruturas da planta. A participação da planta nesta interação é quase que completamente desconhecida, embora se acredite que a planta hospedeira esteja ativamente envolvida no estabelecimento da interação com a bactéria (NOGUEIRA et al. 2001).

2.5. ESTs

Um dos métodos moleculares para estudar a expressão gênica em diferentes tecidos de um organismo eucarionte pluricelular é a análise de cDNA. Sequências parciais de cDNA, também conhecidas como etiquetas de sequências expressas (EST's), podem ser utilizadas para analisar a estrutura, expressão e função de genes (WU et al., 2002). Coleções de ESTs e clones completos de cDNA são importantes para anotação de sequências genômicas e inferência da sua função, refletindo o nível e complexidade de expressão de genes nos tecidos amostrados (KIKUCHI et al., 2003; WU et al., 2002; STERKY et al, 1998). Por exemplo, a análise de uma coleção catalogada de ESTs de arroz forneceu sólido fundamento para auxiliar na caracterização da resposta de defesa da planta contra o fungo *Magnapotha grisea* (JANTASURIYARAT et al. 2005).

Sequenciamento de EST foi o primeiro método usado para rápida identificação de genes expressos (ADAMS et al. 1995). Este método tem sido empregado para identificar os genes que são expressos em vários tecidos, tipos celulares ou estágio de desenvolvimento (OGIHARA et al., 2003; RONNING et al. 2003; MICHALEK et al., 2002), e genes expressos em interação planta-microrganismo (JANTASURIYARAT et al. 2005; FUJIWARA et al., 2004; VENTELON-DEBOUT et al., 2003; NOGUEIRA et al., 2001; KIM et al., 2001; KIM

et al., 2000), como também sob situações de estresse abiótico (HOUDE et al., 2006; SHIMONO et al., 2003; REDDY, et al., 2002)

Numerosos projetos têm o propósito de identificar e caracterizar a coleção completa de genes transcritos (transcriptoma) em organismos alvo (CUI et al., 2004; LAI et al., 2004; LA ROTA e SORRELS, 2004; ZHANG et al., 2004; STERCK et al., 2005). Idealmente, ESTs geradas de uma biblioteca de cDNA total deveriam representar todos os genes expressos no tecido a partir do qual a biblioteca foi construída (REDDY et al. 2002).

A investigação de expressão gênica de plantas durante associações planta-bactéria é uma estratégia para entender os mecanismos moleculares envolvidos neste tipo particular de associação (NOGUEIRA et al., 2001). Foram identificados diversos genes de cana-de-açúcar expressos somente na presença de bactérias diazotróficas endofíticas (NOGUEIRA et al., 2001). O seqüenciamento em larga escala de bibliotecas de cDNA por meio de etiquetas de seqüências expressas (EST) parece ser uma poderosa ferramenta para descobrir novos genes e gerar perfis de expressão gênica de diferentes células e tecidos sob diferentes fases de desenvolvimento e condições fisiológicas (OHLROGGE e BENNING, 2000).

A disponibilização de seqüências de cDNA de arroz tem acelerado a caracterização molecular de genes de interesse e fornece informação de seqüência para planejamento de microarranjo e para anotação do genoma (JANTASURIYARAT et al., 2005). WU et al. (2002) realizaram mapeamento de ESTs para determinar a posição cromossomal de genes expressos em arroz. Neste estudo foi descrito o mapa transcricional de aproximadamente 80% do genoma do arroz. Regiões com alta densidade de ESTs foram encontradas nos cromossomos 1, 2 e 3, contendo 41% do total de ESTs no mapa, a densidade destas regiões é aproximadamente duas vezes maior que a dos cromossomos 11 e 12. A maioria das regiões de ESTs está distribuída nas regiões distais de cada braço dos cromossomos e em menor densidade nos centrômeros. A maioria (40%) das ESTs identificadas ocupa somente 21% de todo o genoma.

2.6 SEQÜENCIAMENTO DO GENOMA DE ARROZ

O genoma de *Oryza sativa* contém aproximadamente 389 milhões de pares de bases (Mb), o menor genoma entre os cereais (MATSUMOTO et al., 2005). O seu seqüenciamento é importante para explorar a possibilidade de melhoramento deste cereal. Foram estimados aproximadamente 38000 a 40000 genes por meio de seqüências de ESTs (YU et. al., 2005). Porém, mais recentemente o RAP-DB (Rice Anotation Project Database) determinou um número menor de genes: as evidências sugerem que 31.439 *loci* são expressos, sendo que 30.192 são potenciais codificadores de proteínas (ITOH, et al., 2007). O conhecimento da localização de todos os genes do genoma aumenta a utilidade de marcadores moleculares permitindo a identificação de genes candidatos envolvidos no controle de uma característica específica de interesse (SASAKI e BURR, 2000). Até recentemente *Arabidopsis thaliana* era a única planta a ter seu genoma seqüenciado, mas as seqüências genômicas de duas culturas de arroz mudaram este panorama. *O. sativa* é atualmente um dos poucos organismos a ter os genomas de duas importantes subespécies (*O. sativa japonica* e *O. sativa indica*) sequenciados completamente (YU et al. 2005). Comparações entre as subespécies *indica* e *japonica* revelaram pequena diferença no conteúdo dos genes: somente 2 a 3% dos genes são únicos para uma das duas subespécies, porém existem enormes diferenças nas regiões intergênicas (YU et al. 2005).

O IRGSP (*The International Genome Sequencing Project*), um consórcio de laboratórios de 10 países iniciaram o seqüenciamento de *O. sativa* ssp. *japonica* em 1998, e em 2004, seqüências 95% do genoma foram disponibilizadas publicamente, sendo que os 5% restantes representam algumas regiões teloméricas e centroméricas (MATSUMOTO et al., 2005; SASAKI e BURR, 2000).

3. OBJETIVOS

3.1 OBJETIVOS GERAIS

Identificar ESTs de raízes de arroz induzidos pela inoculação com *Herbaspirillum seropedicae*.

3.2 OBJETIVOS ESPECÍFICOS

- Sequenciar 2000 clones de bibliotecas de ESTs de raízes de arroz 5 dias após inoculação com *H. seropedicae*;
- Sequenciar 2000 clones de bibliotecas de ESTs de raízes de arroz não inoculado, como controle;
- Anotar as seqüências de ESTs obtidas e determinar as prováveis funções dos produtos de tradução;
- Identificar ESTs presentes preferencialmente em plantas colonizadas por *H. seropedicae*.

4. MATERIAL E MÉTODOS

4.1 BACTÉRIAS E VETOR UTILIZADOS

As características das bactérias e do vetor utilizados para construção das bibliotecas de cDNA estão descritas na Tabela 1 (BRUSAMARELO, 2007).

TABELA 1 – BACTÉRIAS E VETOR UTILIZADO:

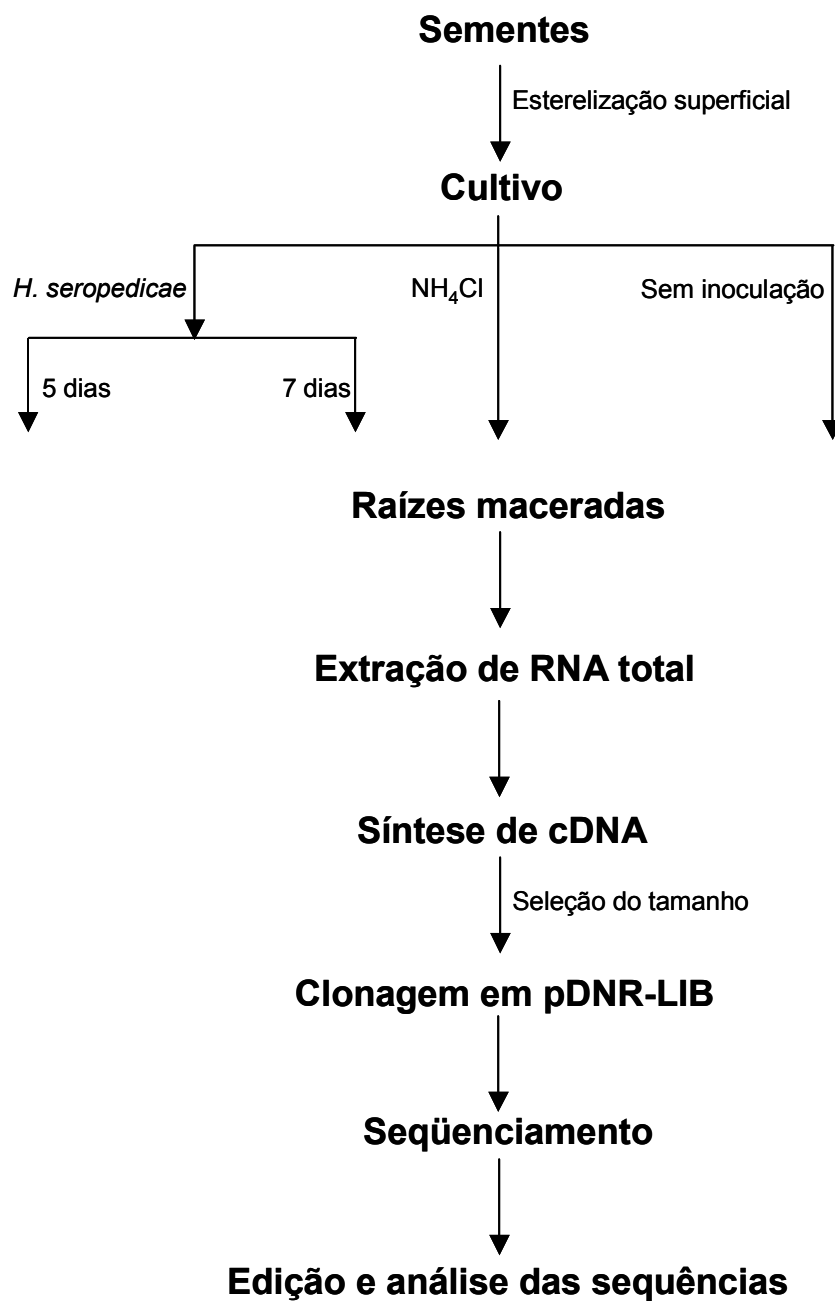
Bactéria	Estirpe	Fenótipo/Genótipo	Referência
<i>Herbaspirillum seropedicae</i>	SMR1	selvagem, Nif ⁺ , Sm ^R	SOUZA et al, 1995
<i>Escherichia coli</i>	TOP10	<i>recA1</i> (Str ^R)	Invitrogen TM
Plasmídeo vetor		Características	
Pdnr-LIB		-Cm ^R	Clontech
		-Clonagem direcionada, mediada recombinase Cre	sítio pela
		-Susceptibilidade a sacarose (SacB)	a

4.2 OBTENÇÃO DE ESTS (BRUSAMARELO, 2007)

As bibliotecas de cDNA de raízes de arroz foram construídas por Brusamarelo (2007). Foi utilizado *Oryza sativa japonica*, cultivar Cateto Zebu, proveniente do Centro Nacional de Pesquisa de Arroz e Feijão (EMBRAPA-CNPAP). Resumidamente, as sementes sofreram esterilização superficial e foram pré-germinadas no escuro por 2-3 dias, as plântulas foram transplantadas para tubos de vidro estéreis contendo 20 mL de meio semi-sólido (0,2% de agar) de Hoagland e então incubadas em câmara de cultivo sob condições de fotoperíodo

14h luz/10h escuro, temperatura de 24°C e umidade superior a 70% por 7 dias. As plantas foram então submetidas a três tratamentos: inoculação com *H. seropedicae* (aproximadamente 10^8 células/mL), adição de NH_4Cl e controle sem inoculação. Após 5 dias de incubação em câmara de cultivo nas mesmas condições anteriores, o RNA total das raízes foi extraído utilizando o kit RNAasy (Quiagen) e utilizado para a síntese de cDNA, que foi clonado em vetor de expressão pDNR-LIB utilizando o *kit Creator SMART cDNA Library Construction* (Clontech). (figura 1)

FIGURA 1 - FLUXOGRAMA DE OBTENÇÃO DE ESTS DE *Oryza sativa*



(Fonte: BRUSAMARELO, 2007)

4.3 MANIPULAÇÃO DE DNA

4.3.1 Extração de DNA plasmidial em placas de 96 poços

A purificação do DNA plasmidial foi realizada por lise alcalina. Colônias transformantes contendo os insertos de cDNA das bibliotecas de cDNA foram crescidas em blocos de 96 poços (*deep well plate*) com 1,25 mL de meio Terrific Broth (Invitrogen) contendo 30 µg/mL de cloranfenicol (Cm) em cada poço e incubado por aproximadamente 16 horas sob agitação, à 37°C. As células foram coletadas por centrifugação (4000rpm, 7 minutos), o precipitado ressuspense em 180 µL de solução GET (Tris-HCl 25 mmol.L⁻¹, pH 8,0; glicose 50 mmol.L⁻¹ e EDTA 10 mmol.L⁻¹). Novamente as células foram centrifugadas a 4000 rpm por 7 minutos e o sobrenadante descartado. Em seguida as células foram ressuspensas sob agitação em 80µL de GET acrescida de 0,125 mg.mL⁻¹ de RnaseA (Invitrogen) e transferidas para microplaca de 96 poços com capacidade para 250 µL. As células foram lisadas com 80 µL de solução de lise (NaOH 0,2M e SDS 1%) homogeneizadas por inversão. As proteínas, o DNA cromossomal desnaturado, os resíduos celulares e o SDS foram precipitados com a adição de 80 µL de acetato de potássio 3 mol.L⁻¹ pH 5,5 gelado. Em seguida o lisado foi incubado por 10 minutos em temperatura ambiente, por 30 minutos à 90°C e finalmente 15 minutos no gelo. Após centrifugação a 4000rpm por 10 minutos sob temperatura de 4°C, o sobrenadante foi coletado (~130µL) e transferido para placa de filtro Millipore (MAGV N22) contendo microplaca de coleta acoplada. As amostras foram centrifugadas por 5 minutos à 4000rpm sob temperatura de 20°C. Ao filtrado foi acrescentado 0,6 volumes de isopropanol, centrifugado por 45 minutos (4000rpm, e 8°C) para precipitação do DNA plasmidial, que foi lavado em seguida com 150µL de etanol 70%, seco à temperatura ambiente por aproximadamente 12 horas e dissolvido em 30 µL de água ultra-pura estéril. Para visualização do DNA obtido, amostras foram submetidas a eletroforese em gel de agarose 1%, coradas com brometo de etídeo e visualizadas em transluminador de ultravioleta.

4.3.2 Eletroforese de DNA

O perfil eletroforético dos plasmídeos foi determinado em gel de agarose (SAMBROOK et al., 1989). O gel de agarose foi preparado numa concentração de 1% em tampão TBE 1X (SAMBROOK et al., 1989). A corrida eletroforética foi realizada durante 1 hora a 100 V. Para visualização do DNA, o gel de agarose foi corado com brometo de etídeo (0,5 µg/mL), pois este composto intercala-se ao DNA, formando um complexo fluorescente que pode ser visualizado sob luz ultravioleta.

4.3.3 Reação de Sequenciamento

O seqüenciamento de DNA plasmidial foi realizado pelo método de terminação de cadeia por incorporação de di-desoxinucleotídeo marcado com fluoróforo (KARGER, HARRIS e GESTELANG, 1991). No sistema da reação foi utilizado 100 a 400 ng de DNA purificado, 8 pmol do oligonucleotídeo iniciador universal (GACGTTGTAAAACGACGGCCAGT), 3µL de reativo ET Terminator mix (GE Health Care) e água ultra-pura estéril suficiente para 10µL, em uma placa de 96 poços. O oligonucleotídeo universal anela-se ao vetor na extremidade correspondente à 5' do inserto. A reação de sequenciamento foi conduzida em termociclador com as seguintes condições: 1 ciclo de 1 minuto a 95°C, seguido por 35 ciclos de 94°C por 20 segundos, e 60°C por 1 minuto e 30 segundos. Os produtos da reação foram purificados por precipitação com 2 µL de acetato de amônio (7,5 M) e três volumes de etanol absoluto, incubado por 20 minutos em freezer, em seguida centrifugado à 4000 rpm sob 8°C por 45 minutos. O sobrenadante foi descartado e o precipitado lavado com 150 µL de etanol 80%, novamente centrifugado por 15 minutos. O DNA foi seco em temperatura ambiente pelo período de aproximadamente 12 horas. Os produtos de reação foram analisados em um seqüenciador automático de DNA, MegaBace1000 (GE Health Care).

4.4 ANÁLISE DE BIOINFORMÁTICA DAS ESTs

A análise inicial das seqüências foi realizada por meio de *scripts* desenvolvidos por BRUSAMARELO (2007) em linguagem BASH, que é um interpretador de comandos entre o sistema operacional e o usuário. Os *scripts* possibilitaram a automatização da criação de diretórios, nomeação dos arquivos de seqüência e agrupamento de seqüências contíguas, sendo que as etapas de busca por homologia em bancos de dados e análise dos resultados foram realizadas manualmente. Na figuras 2 e 3 são mostrados fluxogramas da análise das ESTs.

FIGURA 2 – FLUXOGRAMA GERAL DA ANÁLISE DAS ESTS IN SILICO:

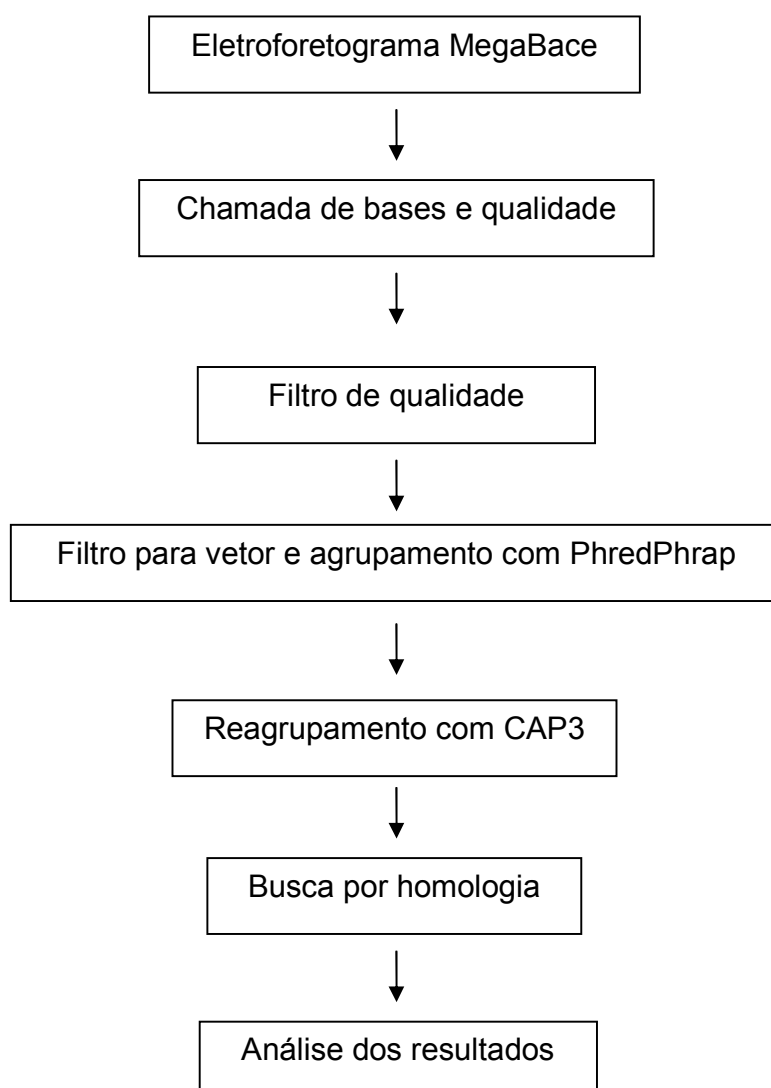
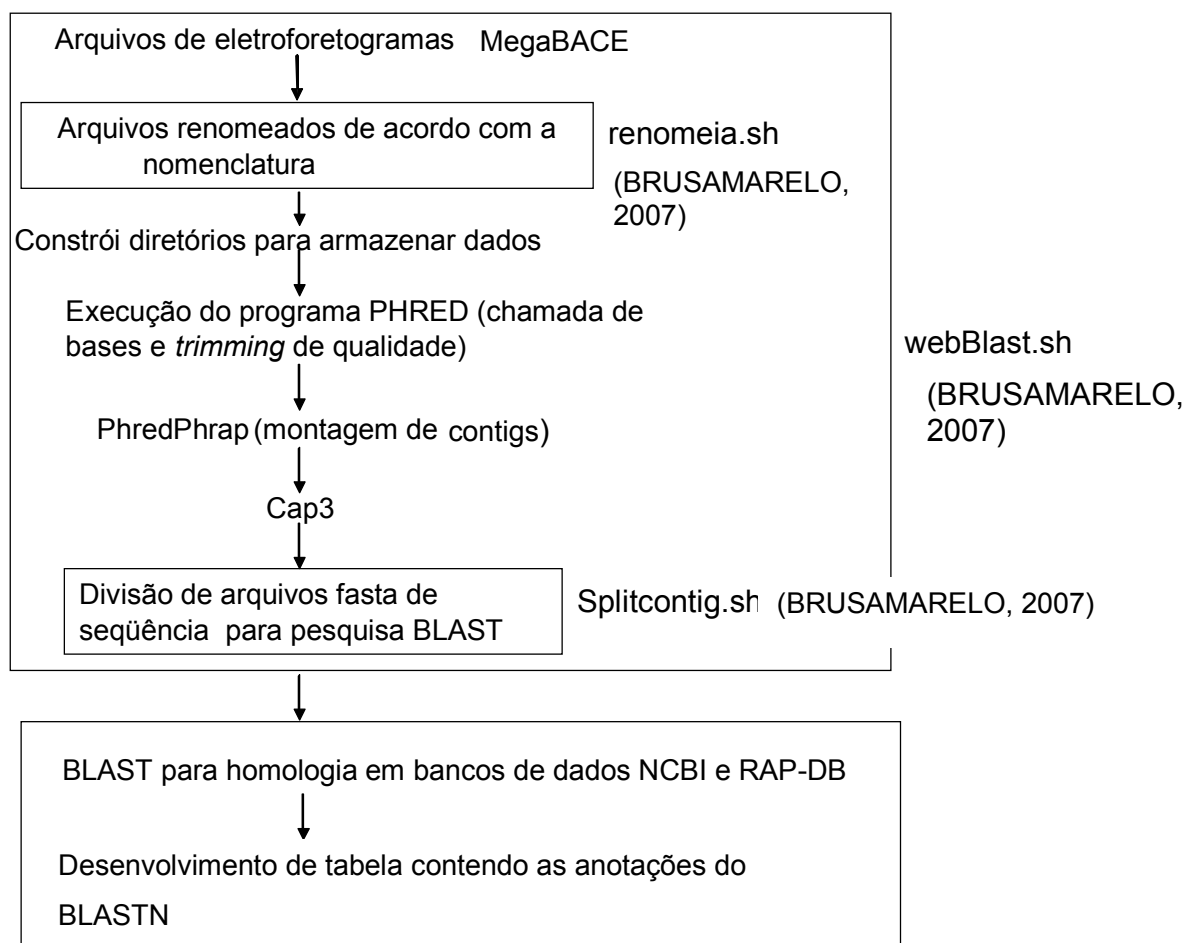


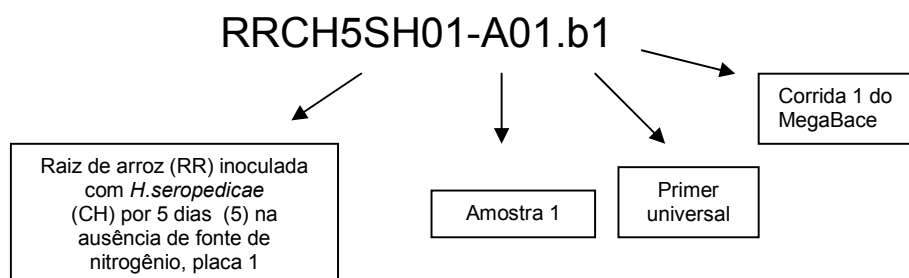
FIGURA 3 – FLUXOGRAMA DETALHADO DA ANÁLISE DAS ESTS



4.4.1 Nomenclatura utilizada para nomear as seqüências de ESTs

Os eletroforetogramas gerados pelo MegaBace1000e foram nomeados de acordo com a biblioteca da qual foram obtidas as seqüências (RRSHSN ou RRCH5SN), seguida do número da placa, da posição na placa, do oligonucleotideo iniciador utilizado (b: universal) e o número da corrida do MegaBace1000 (figura 4).

FIGURA 4 – NOMENCLATURA UTILIZADA PARA OS ELETROFORETOGRAMAS



4.4.2 Edição, montagem e alinhamento das seqüências.

O primeiro *script* renomeia.sh (Brusamarelo, 2007) renomeia os eletroforetogramas gerados pelo Megabace1000. São construídos diretórios para armazenagem destes arquivos. As seqüências são editadas pelo programa Phred (EWING et al, 1998; EWING e GREEN, 1998), que determina a identidade das bases, a qualidade (probabilidade de que a identificação da base seja incorreta) e elimina ou filtra as seqüências do vetor. Posteriormente, o programa PhredPhrap são executados para montagem das seqüências contíguas. Os contigs gerados são remontados pelo programa CAP3 (HUANG e MADAN, 1999). Este programa foi executado com as opções -o 50 -p 90, que determina que um mínimo de 50 pb de sobreposição e identidade de pelo menos 90% para efetuar alinhamento entre seqüências contíguas. O *script* Splitcontig.sh divide as seqüências geradas em arquivos denominados contigs ou singlets. Contigs representam seqüências contíguas alinhadas a partir das leituras, enquanto singlets representam leituras que não encontraram sobreposição com outras para formação de seqüência contígua. O script webBlast.sh integra os *scripts* anteriores. As seqüências finais são fornecidas em no formato FASTA para análise pelo programa BLAST. As

funções e referencias dos *scripts* e programas de bioinformática utilizados para montagem e análise das seqüências estão mostradas nas tabelas 2 e 3.

TABELA 2 – SCRIPTS UTILIZADOS PARA PROCESSAMENTO DOS ELETROFORETOGRAMAS

Script	Referência	Funções
renomeia.sh	Brusamarelo, 2007.	Renomear os eletroforetogramas gerado pelo seqüenciador Megabase1000
Splitfastfile.sh	Brusamarelo, 2007.	Dividir os arquivos fasta dos " <i>contigs</i> " e " <i>singlets</i> " em arquivos Fasta separados
WebBlast.sh	Brusamarelo, 2007.	Integrar a execução de todos os programas e scripts desde renomear diretórios a execução do Cap3

TABELA 3 – PROGRAMAS DE BIOINFORMÁTICA UTILIZADOS NA MONTAGEM E ANÁLISE DAS SEQUÊNCIAS DE EST

Programa de bioinformática	Atribuições	Referência
Phred	Recebe os arquivos de eletroforetogramas, faz a identificação de bases e atribui valor de qualidade para cada base.	EWING et al., 1998 EWING, GREEN, 1998
Phrap	Realiza a montagem das seqüências contíguas a partir de regiões de sobreposição	http://www.phrap.org/phredphrap/phrap.html
CAP3	Realiza remontagem das seqüências contíguas	HUANG, MADAN, 1999
BLAST	Compara seqüências contra um banco de dados	ALTSCHUL et al., 1997

4.4.3 Busca por homologia em bancos de dados

As seqüências parciais de cDNA foram comparadas com seqüências do banco de seqüências público NCBI (www.ncbi.nlm.nih.gov) e RAP-DB (Rice Annotation Project - Data Base) (<http://rapdb.dna.affrc.go.jp/>). Para a pesquisa no banco de dados GenBank nr, foi utilizado o programa MEGABLAST, que compara várias seqüências simultaneamente, contidas em um único arquivo, contra um banco de dados de nucleotídeos de todos os organismos que possuem seus genomas anotados. Também foi utilizado o programa BlastX contra o banco nr (non-redundant protein sequences) quando não foi identificada similaridade com banco de seqüências de DNA genômico e EST. Para análises complementares foi utilizado o banco de dados de transcritos e genes preditos do RAP-DB, que

apresenta o *loci* RAP (conjunto de genes anotados) com o correspondente *locus* IDs (predição do gene anotado) representando os genes anotados. Sequências de cDNA do “The Institute for Genomic Research” (TIGR) também foram incorporadas ao RAP-DB. Cada sequência de cDNA do RAP-DB possui ligação eletrônica com o portal *Gene Ontology* (www.geneontology.org), que contém informações sobre estrutura, função e localização celular do provável gene (OHYANAGI et al, 2006).

5. RESULTADOS E DISCUSSÃO

Neste trabalho duas bibliotecas de cDNA de raízes de arroz foram seqüenciadas e comparadas. Uma das bibliotecas foi inoculada com *H.seropediacae* (raízes coletadas 5 dias após inoculação) e a outra não sofreu tratamento e foi usada como controle. Esta comparação foi destinada a identificar genes que tivessem sua expressão diferenciada na presença deste diazotrofo, permitindo assim uma melhor compreensão dos fatores moleculares envolvidos na interação planta-bactéria.

As extremidades 5' de 4608 clones das bibliotecas RRCH5SN e RRSN foram seqüenciadas. Os eletroforetogramas obtidos foram processados pelo programa Phred, que identifica bases e atribui um valor de confiabilidade para cada base. Neste trabalho foi definido que as seqüências obtidas deveriam ter um comprimento mínimo de 100 pares de base com valor de qualidade Phred 15, que seria a probabilidade de 3 erros a cada 100 pares de base. Das 4608 reações de seqüenciamento realizadas, 2314 resultaram em seqüências com tamanho e qualidade definido acima (tabela 4).

TABELA 4 - CARACTERÍSTICAS DAS BIBLIOTECAS DE cDNA DE RAÍZES DE ARROZ UTILIZADAS

Bibliotecas de ESTs de arroz	Nº clones estocados	Nº seqüências com iniciador universal	Nº de leituras de qualidade obtidas	Nº singlets	Nº de contigs	ESTs únicas
RRCH5SN*	2976	2496	1303	701	54	755
RRSN*	2784	2112	1011	560	62	622
TOTAL	5760	4608	2314	1261	116	1377

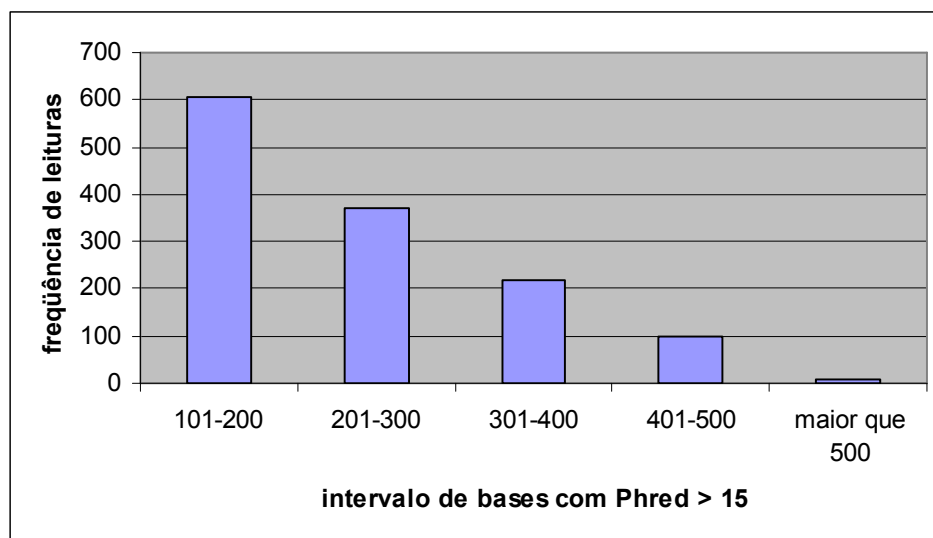
*Bibliotecas BRUSAMARELO, 2007

5.1 SEQUENCIAMENTO

O seqüenciamento de clones das bibliotecas RRCH5SN (inoculada com *H. seropedicae* por 5 dias) e RRSN (controle sem inoculação) foi realizado utilizando-se o oligonucleotídeo iniciador universal M13 (VIEIRA e MESSING, 1982), que permite seqüenciar a extremidade 5' do cDNA.

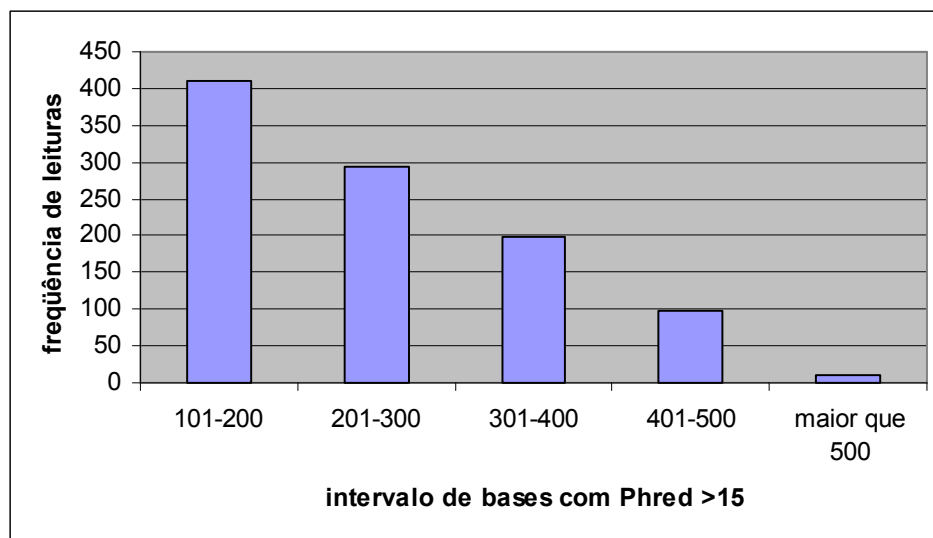
Foram seqüenciados 4608 clones (tabela 4), porém nem todas as seqüências obtidas tinham os valores mínimos de qualidade (100 pares de base, com Phred acima de 15). Foram obtidas 2314 seqüências, que foram utilizadas na montagem de regiões contíguas utilizando o programa Phrap e posteriormente o programa CAP3. A biblioteca RRCH5SN apresentou 1303 leituras acima de 100 pares de base e Phred maior que 15 (gráfico 1), destas foram obtidas 701 seqüências únicas (singlets) e 54 seqüências contíguas (cada uma contendo 2 a 27 seqüências individuais), totalizando 755 ESTs únicas com identidade acima de 90% com os bancos de dados NCBI, que contém seqüências anotadas de organismos modelo de diversos reinos e/ou RAP-DB, que contém seqüências do genoma do arroz, além de prováveis transcritos e genes preditos deste organismo especificamente. Para a biblioteca RRSN foram obtidas 1011 leituras (gráfico 2), das quais 560 representaram seqüências únicas e 62 seqüências contíguas (cada uma com 2 a 16 seqüências), num total de 622 ESTs únicas, também com identidade acima de 90%.

GRÁFICO 1 – DISTRIBUIÇÃO DAS SEQÜÊNCIAS DE ESTs DA BIBLIOTECA RRCH5SN DE ACORDO COM SEU TAMANHO E QUALIDADE



As regiões de baixa qualidade das extremidades das leituras foram removidas anteriormente por meio do filtro de qualidade a partir do programa Phred. Cada coluna representa o número de leituras contendo os números de bases indicados com Phred maior que 15. A biblioteca RRCH5SN apresentou 1303 leituras com 100 ou mais pares de base com qualidade Phred 15 ou superior.

GRÁFICO 2 – DISTRIBUIÇÃO DAS SEQÜÊNCIAS DE ESTs DA BIBLIOTECA RRSN DE ACORDO COM SEU TAMANHO E QUALIDADE



As regiões de baixa qualidade das extremidades das leituras foram removidas anteriormente por meio do filtro de qualidade a partir do programa phred 15 ou superior. Cada coluna representa o número de leituras contendo os números de bases indicados com Phred maior que 15. A biblioteca RRSN apresentou 1011 leituras com 100 ou mais pares de base com qualidade phred 15 ou superior.

5.2 BUSCA POR SIMILARIDADE EM BANCOS DE DADOS

As seqüências obtidas foram então utilizadas para pesquisa de similaridade com seqüências com bancos de dados NCBI (<http://www.ncbi.nlm.nih.gov>) e RAP-DB (<http://rapdb.dna.affrc.go.jp/>) utilizando o programa BLAST que realiza alinhamento local das seqüências. Primeiramente foi utilizado o programa BLASTN para comparação com bancos de dados de todos os organismos anotados no banco NCBI e o banco de seqüências de arroz RAP-DB. As seqüências de ESTs que não possuíam descrição como provável gene ou cDNA foram analisados com o programa BLASTX. O resultados desta busca estão mostrados nos anexos depositados na coordenação do programa de Pós-Graduação em Bioquímica..

Das 755 ESTs únicas da biblioteca RRCH5SN, 82,5% das seqüências possuíam identidade $\geq 90\%$ com genes de arroz, 16,2% apresentam identidade com DNA genômico de arroz e 1,3% correspondem a cDNAs de arroz não caracterizados.

Das 622 ESTs únicas da biblioteca RRSN, com as qualidades mínimas estabelecidas para este trabalho, 72% correspondem a genes de arroz, 27,5% correspondem a DNA genômico com identidade $\geq 90\%$ com os bancos de dados e 0,5% correspondem a cDNAs de arroz não caracterizados.

As 293 seqüências que apresentaram esta identidade apenas contra DNA genômico de arroz foram reanalisadas utilizando o programa BLASTX que comparou com o banco de proteínas do NCBI. Posteriormente 10 seqüências de cada biblioteca aleatoriamente escolhidas, foram alinhadas com o genoma do arroz usando o programa BLASTN e o banco RAP-DB. Destas, 30% das seqüências idênticas a DNA genômico, alinharam com regiões 5' não traduzidas. Além disso, uma vez que foi utilizado RNA total para construir as bibliotecas é possível que o DNA cromossomal contaminante tenha sido clonado. Finalmente estas seqüências podem ainda representar novos genes de arroz. A comparação com o banco de dados de cDNA de outros organismos (ex. *Arabidopsis thaliana*) poderá ajudar a esclarecer a origem destes clones.

Entre as seqüências da biblioteca RRSN, 38,9% correspondem a genes contidos nos cromossomos 1, 2 ou 3 (gráfico 3). Na biblioteca RRCH5SN, os mesmos cromossomos contêm 37,1% dos genes identificados (gráfico4). Estes resultados são semelhantes aos de WU et al. (2002) que reportaram que 40% de ESTs de arroz, correspondem a genes dos cromossomos 1, 2 ou 3 por meio de mapeamento de ESTs. Segundo WU et al. (2002) 40% dos genes de arroz está contido em 21% do genoma.

GRÁFICO 3 - DISTRIBUIÇÃO DE SEQUÊNCIAS DA BIBLIOTECA RRSHSN DE cDNA DE RAÍZES DE ARROZ NOS REFERENTES CROMOSSOMOS DE *Oryza sativa*

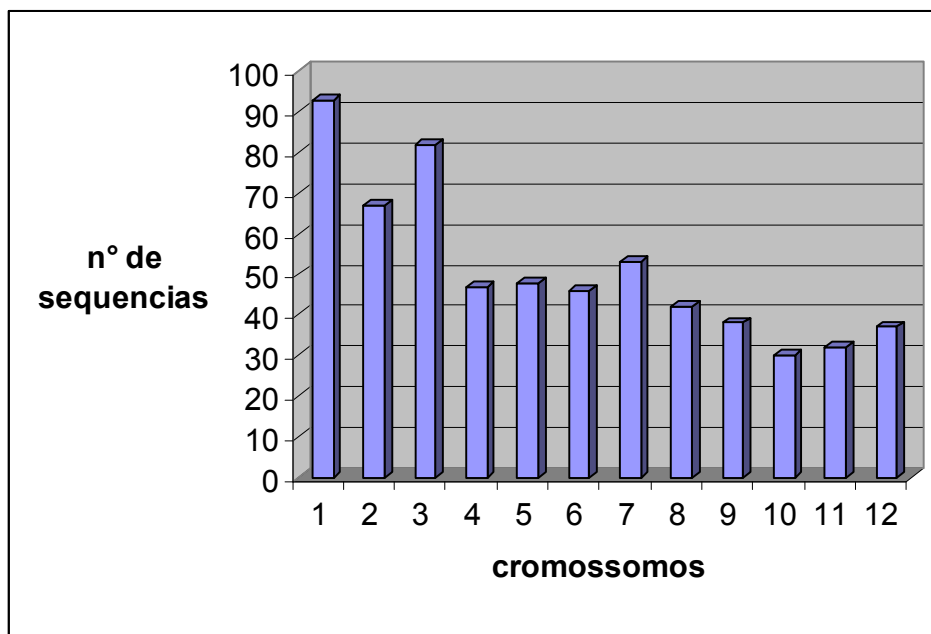
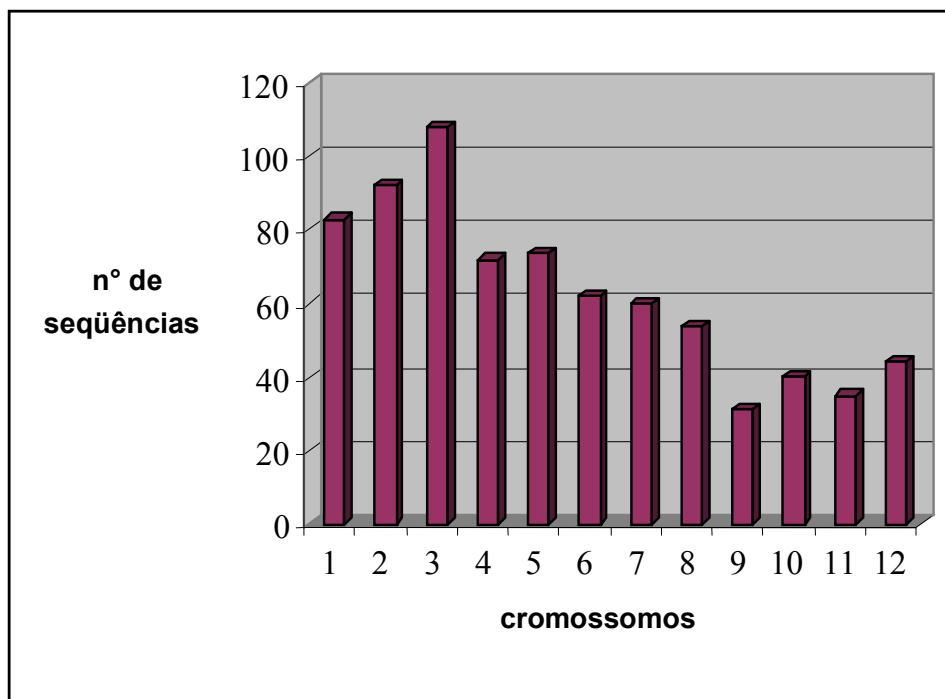


GRÁFICO 4 - DISTRIBUIÇÃO DE SEQÜÊNCIAS DA BIBLIOTECA RRCH5SN DE cDNA DE RAÍZES DE ARROZ NOS REFERENTES CROMOSSOMOS DE *Oryza sativa*

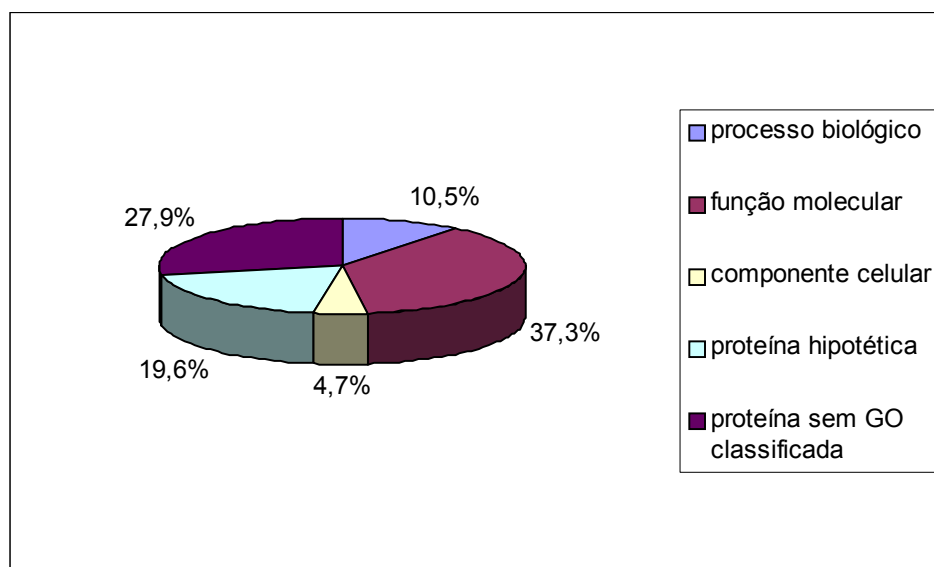


5.2.1 Categorização funcional

Os genes identificados neste trabalho foram categorizados funcionalmente segundo suas GOs. O projeto “Gene Ontology” (GO) (www.geneontology.org) representa um esforço para definir uma descrição consistente de produtos gênicos e suas funções. Para isto, as seguintes categorias foram definidas: função molecular, processo biológico e componente celular. O termo GO consiste de um identificador único na forma GO: nnnnnnn, contendo um nome comum, sinônimos e uma definição para cada termo.

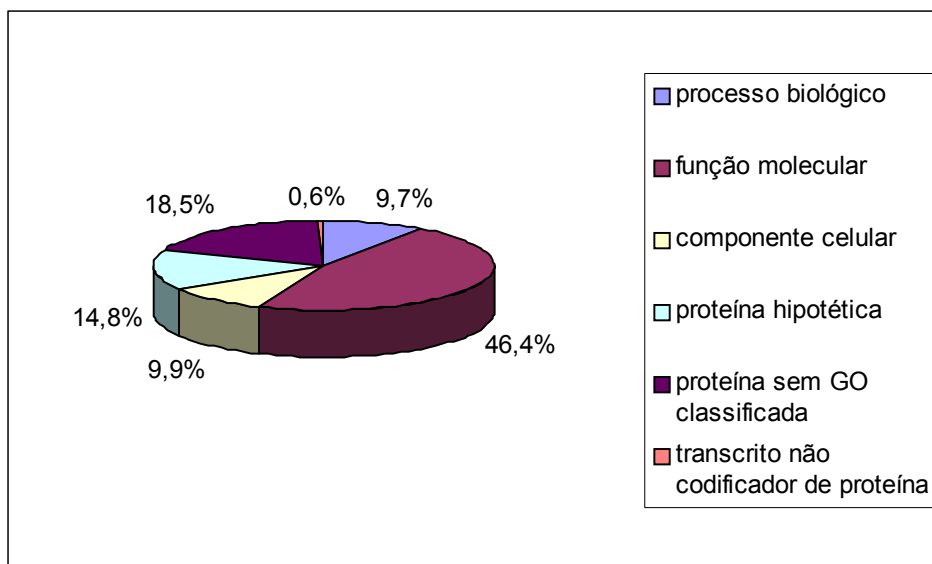
Entre as seqüências da biblioteca controle 10,5% correspondem a genes que pertencem à categoria processo biológico, 37,3% à função molecular, 4,7% à componente celular, 19,6% são proteínas hipotéticas de arroz e 27,9% das ESTs do arroz não foram classificadas dentro da terminologia GO (gráfico 5).

GRÁFICO 5 – CLASSIFICAÇÃO FUNCIONAL DA BIBLIOTECA RRSNHSN BASEADA NO GENE ONTOLOGY.



A distribuição das seqüências da biblioteca RRCH5SN entre categorias funcionais foi a seguinte: 9,7% pertencem a processo biológico, 46,4% a função molecular, 9,9% componente celular, 14,8% eram proteínas hipotéticas de arroz, 18,5% proteínas de arroz com GOs não encontradas e 0,6% eram transcritos não codificadores de proteínas (gráfico 6)

GRÁFICO 6 – CLASSIFICAÇÃO FUNCIONAL DA BIBLIOTECA RRCH5SN BASEADA NO GENE ONTOLOGY.



5.3 IDENTIFICAÇÃO DOS GENES PROVAVELMENTE EXPRESSOS DIFERENCIALMENTE

A partir da comparação das bibliotecas estudadas neste trabalho, foram selecionados genes que foram expressos apenas na biblioteca inoculada com *H. seropedicae* ou que foram expressos em maior número na presença deste diazotrofo.

5.3.1 Genes encontrados exclusivamente na biblioteca inoculada com *H. seropedicae*

A partir de uma comparação realizada entre as bibliotecas RRSN e RRCH5SN, foram identificados e selecionados genes que possuem expressão apenas na biblioteca RRCH5SN (tabela 5).

Nesta análise para verificação de genes expressos apenas na biblioteca inoculada foram encontradas 3 genes codificando proteínas contendo domínio WD40, dois genes Os05g0405900 e um Os04g0481600. Estas prováveis proteínas codificadas contendo o domínio WD40 são caracterizadas por seu envolvimento na regulação da transcrição da biossíntese de flavonóides em *A. thaliana* e outras espécies (BROUN, 2005). Flavonóides constituem uma classe diversa de componentes naturais produzidos como resultado do metabolismo secundário da planta (SHAW et al., 2006). Em leguminosas, flavonóides são sinais chave na iniciação da formação do nódulo na simbiose de fixação de nitrogênio através da ativação de agentes de indução de nodulação de Rhizobia e genes relacionados à nodulação (BROUGHTON et al., 2000).

Um gene codificando uma provável proteína SEP 2 (*stress enhanced protein 2*) (Os04g0639200) foi encontrado unicamente também na biblioteca inoculada. ESTs de SEP2 foram encontradas em *A.thaliana* sob condições de estresse luminoso (HEDDAD e ADAMSKA, 2000).

Foram encontrados dois prováveis transportadores de nitrato Os01g0704100 e Os03g0193300. Transportadores de nitrato foram encontrados exclusivamente quando cana-de-açúcar foi inoculada com os diazotrofos *Gluconacetobacter diazotrophicus* e *Herbaspirillum rubrisubalbicans*. NOGUEIRA et al. (2001) sugere que a planta talvez expresse estes transportadores porque esteja assimilando nitrogênio fixado suprido pela bactéria.

Uma única sequência codificando uma provável lipoxigenase 2 foi encontrada somente na biblioteca inoculada. Lipoxigenases catalisam hidroperoxidação de ácidos graxos. O produto de reação, ácidos graxos

hidroperóxidos, são então metabolizados em componentes fisiologicamente ativos (OHTA et al., 1992). O cDNA de lipoxigenase L-2 de arroz foi clonado e foi caracterizado por OHTA et al. (1992). Biossíntese de componentes anti-fungo por meio da via de lipoxigenases foi proposto por OHTA et al. (1990) em sementes de arroz. Os mesmos autores também demonstraram um aumento na atividade de lipoxigenase em folhas de arroz em resposta a infecção com fungo *Magnaporthe grisea* (OHTA et al., 1991). Também foi observado aumento da atividade de lipoxigenases em cana-de-açúcar quando inoculada *Gluconacetobacter diazotrophicus* e *Herbaspirillum rubrisubalbicans*, NOGUEIRA et al. (2001). Aumento da atividade de lipoxigenase é comumente observado em interações planta-bactéria (FINDANTISEF et al., 1999).

Alguns dos genes expressos exclusivamente na biblioteca inoculada podem permitir a inferência de que sua expressão é aumentada na interação *H.seropedicae*-arroz. Estes genes terão sua expressão diferencial confirmada por experimentos futuros, o que permitirá a confirmação de novos genes relacionados com a interação planta-bactéria.

TABELA 5 – GENES REPRESENTADOS APENAS NA BIBLIOTECA RRCH5SN

Clone	Id	Comprimento da sequência (bp)	Número de bases Alinhadas (bp)	Descrição	Cromossomo	Número de acesso	GO_id	Ontologia
RRCH5SN03-D12.b1	93%	198	153/163	Provável transportador de nitrato	1	J100043G11 Os01g0704100 LOC_Os01g50820	GO:0005215: atividade transportadora	Função molecular
RRCH5SN09-F03.b1	99%	390	386/387	Provável lipoxigenase 2 (EC 1.13.11.12 Lipoxigenase L-2)	3	J100040C15 Os03g0738600 LOC_Os03g52860	GO:0003824 : atividade catalítica	Função molecular
RRCH5SN16-D09.b1	100%	295	291/292	Provável SEP 2 (proteína ativada por estresse)	4	J075064G13 Os04g0639200 LOC_Os04g54630	Não encontrada	
RRCH5SN19-D10	98%	292	287/290	Proteína contendo domínio tipo WD40	4	GENE ID: 4336188 Os04g0481600 NP_001053112.1	GO:0004252 : atividade endopeptidase tipo serina	Função molecular
RRCH5SN22-A11	98%	176	171/173	Proteína contendo domínio tipo WD40	5	GENE ID: 4338754 Os05g0405900 NP_001055510.1	GO:0005739 : Mitocôndria	Componente celular
RRCH5SN24-B02	98%	415	405/412	Proteína contendo domínio tipo WD40	5	GENE ID: 4338754 Os05g0405900 NP_001055510.1	GO:0005739 : mitocôndria	Componente celular
RRCH5SN26-H10.b1	98%	108	101/103	Similar a transportador de nitrato	3	J013134C19 Os03g0193300	Não encontrada	

Id = identidade com o banco de dados

Go_Id: número de acesso GO do banco de dados *Gene Ontology*

5.3.2 Genes encontrados preferencialmente na biblioteca inoculada com *H. seropedicae*

A comparação de cDNA das bibliotecas de raízes de arroz revela que alguns dos transcritos de resposta a fitormônios, como etileno e auxina, apresentaram-se em maior número na biblioteca das plantas inoculadas com *H. seropedicae*.

Na biblioteca inoculada com *H. seropedicae* foram encontrados 3 genes de resposta a etileno, enquanto na biblioteca controle, foi encontrado apenas 1 gene relacionado a este tipo de resposta. Três sequências de genes relacionados a resposta a auxina foram encontrados na biblioteca inoculada, e na biblioteca controle, apenas um (tabelas 6 e 7).

5.3.2.1 Genes de resposta ao etileno encontrados preferencialmente na biblioteca inoculada com *H. seropedicae*

Dos genes presentes em maior número na biblioteca inoculada com *H. seropedicae*, 3 estão relacionados com resposta ao etileno: o primeiro codifica uma metionina sintase responsiva a etileno (Os12g0624000). O segundo gene (Os03g0183300) codifica uma provável proteína que contém o domínio AP2, capaz de ligar o elemento responsivo a etileno. Finalmente, foi identificado também um gene (Os07g0617000) codificando uma proteína tipo EREB, um provável elemento capaz de interagir com o fator 5 responsivo a etileno (ERF5) (FUJIMOTO et al., 2000).

Na biblioteca inoculada com *H. seropedicae* o gene Os12g0624000 apareceu duas vezes e codifica uma metionina sintase responsiva a etileno que é similar à proteína Q9SWV3 codificada pelo gene responsivo a etileno ER69, identificado em tomateiro por *differential display* durante os estágios iniciais de maturação dos frutos (ZEGZOUTI et al., 1999). Foi encontrado um representante deste mesmo gene na biblioteca de cDNA de raízes não inoculadas.

Outro gene de resposta a etileno é um fator transcricional relacionado à patogênese, Os03g0183300, que codifica uma provável proteína ligadora a elemento responsivo a etileno contendo o domínio AP2. Este é um domínio de ligação a DNA encontrado em reguladores transcricionais de planta como APELA2 e EREBP (proteína de ligação ao elemento responsivo ao etileno). Em EREBPs o domínio liga especificamente à sequência consenso AGCCGCC (“GCC Box”) de elemento de resposta a etileno (ERE). EREBPs envolvidos em resposta de estresse foram encontrados após clonagem e caracterização de cDNAs correspondentes a EREBPs de *Nicotiana* spp (OHME-TAKAGI e SHINSHI, 1995).

O outro gene identificado na biblioteca de cDNA de plantas inoculadas também codifica uma proteína de resposta a etileno, pertencente a uma classe de EREBP, e conhecido como ERF (*ethylene response factor*) (FUJIMOTO et al., 2000), está no *locus* Os07g0617000 e codifica um provável elemento ligante ao fator 5 responsivo a etileno (ERF5). Proteínas ERF ligam-se especificamente ao “GCC Box”, e foram isoladas de folhas de tabaco (OHME-TAKAGI e SHINSHI, 1995). Foram isolados cDNAs codificando proteínas ERF em *A. thaliana* (AtERF5) por hibridização com cDNA de ERF de tabaco, e foi investigada sua função na regulação transcricional em plantas (FUJIMOTO et al., 2000). FUJIMOTO et al. (2000) demonstrou que ERF5 atua como ativador transcricional de genes dependentes de “GCC Box”. ERF5 possui um provável sítio de fosforilação por MAP quinase conservado na região C-terminal. A expressão de AtERF5 é regulada positivamente sob condições de estresse de frio em *Arabidopsis*. Por sua vez, AtERF5 são ativadores transcricionais, que respondem a sinais extracelulares e modulam a expressão de genes mediada por “GCC box” (FUJIMOTO et al., 2000).

Etileno é um hormônio vegetal gasoso que afeta vários processos de desenvolvimento e respostas adaptativas, incluindo germinação, senescência de flores e folhas, amadurecimento de frutos, abscisão da folha, morte celular programada, resposta a estresse e ataque de patógenos. O controle destes processos por etileno envolve regulação de biossíntese e percepção seguida por sinalização para fatores de transcrição que regulam genes de resposta a etileno (OHME-TAKAGI, SUSUKI e SHINSHI, 2000).

Cavalcante et al. (2007) mostraram a regulação diferencial de genes codificando membros da via de sinalização de etileno durante associação entre cana-de-açúcar e bactéria *Herbaspirillum* spp e *G. diazotrophicus*. Estes autores caracterizaram pela primeira vez um provável receptor de etileno ER1 e dois prováveis fatores responsivos a etileno ERF1 e ERF2 de cana-de-açúcar. O padrão de expressão destes genes em resposta à associação com bactérias diazotróficas endofíticas e uma bactéria patogênica *Leifsonia xyli*, foi investigado em diferentes variedades de cana-de-açúcar como: SP70-1143, Chunee e SP70-3370. AO padrão expressão de ER1, ERF1 e ERF2 em resposta aos microrganismos benéficos e patogênicos sugere que estes genes tenham função na sinalização de defesa da planta (CAVALCANTE et al., 2007).

O fitormônio etileno também participa na sinalização de interações planta-bactéria durante a formação do nódulo na leguminosa *Medicago trunculata* quando inoculada com *Rhizobium meliloti* (PENMETSA e COOK, 1997) e na expressão de genes relacionados à patogênese (OHME-TAKAGI, SUSUKI e SHINSHI, 2000).

Cao et al. (2006) encontraram que a expressão de genes ERF de arroz foi induzida pela infecção com o fungo *M. grisea*, e mostraram que a ativação de alguns genes que codificam proteínas ERF, induzida por patógeno, pode estar envolvida na mediação de resistência a doenças de arroz. Estresse abiótico (salino, sal e seca) também induz a expressão destes genes (CAO et al., 2006). As ESTs de OsEREBP1 e OsARF foram induzidas quando arroz foi inoculado como fungo *Magnaporthe grisea*, e supõe-se sua expressão seja ativada pela infecção com o patógeno (KIM et al., 2000).

5.3.2.2 Genes de resposta a auxina encontrados preferencialmente na biblioteca inoculada com *H.seropedicae*.

Foram também identificados nas bibliotecas de cDNA de arroz genes que codificam proteínas que respondem à auxina: 3 na biblioteca RRCH5SN (Os02g0228900, Os05g0559400 e Os01g0670800), e 1 na biblioteca RRSN (Os01g0927600).

O gene Os02g0228900 codifica uma proteína responsiva a auxina, denominada IAA7 (proteína 7 induzida por ácido indol acético) pertencente à família Aux/IAA e que contém o domínio Aux-IAA-ARF.

O produto do gene Os05g0559400 é uma proteína da família Aux/IAA, denominada OsIAA19, responsiva a auxina (proteína 19 induzida por ácido indol acético) e que contém o domínio Aux-IAA-ARF.

Proteínas Aux/IAA são fatores transcricionais de vida curta que funcionam como repressores de genes de resposta inicial a auxina sob baixas concentrações de auxina. Análises de mutantes de genes Aux/IAA demonstraram que têm função central na regulação do crescimento e desenvolvimento em *A. thaliana* (TIAN e REED, 1999; ROUSE et al., 1998).

Os01g0670800 é um gene que codifica um fator 2 responsivo a auxina (ARF 2) ARFB. Existem os seguintes nomes alternativos para esta proteína: ETTIN2 (OsETTIN2) e ARFs ("Auxin response factors"). Proteínas da família ARF contêm domínio TF-B3 (fator transcricional B3) de ligação ao DNA.

Na biblioteca não inoculada (RRSN) também foi encontrado um gene (Os01g0927600) que codifica uma proteína da família Aux/IAA. Esta proteína é uma ARF4 (fator de resposta à auxina)/ ARFD, também denominada OsARF2. Esta proteína é um dímero que contém os domínios Aux/IAA- ARF e TF-B3 de ligação a DNA, e é expressa em raízes, caules, folhas e panículas jovens (WANG et al., 2007).

ARFs são fatores transcricionais que ligam especificamente à sequência de DNA 5'-TGTCTC-3' encontrado em elementos promotores responsivos à auxina (AuxREs) (ULMASOV et al., 1999).

As famílias de proteínas Aux/IAA e ARF (*Auxin Response Factors*) são reguladores chave na expressão de genes modulados por auxina. Ambas proteínas funcionam como reguladores transcricionais (LISCUM e REED, 2002; HAGEN e GUILFOYLE, 2002). Proteínas ARF ligam-se a promotores responsivos a auxina (AuxREs) usando um domínio amino terminal de ligação a DNA. ARF pode atuar como ativador ou repressor transcricional dependendo da sequência da região interna da proteína (TIWARI, HAGEN e GUILFOYLE, 2003). Proteínas Aux/IAA provavelmente regulam a transcrição modificando a atividade de ARF. Ambas famílias de proteínas possuem em comum um domínio C-terminal de interação proteína-proteína. Este domínio está envolvido em homo e heterodimerização entre os membros das proteínas ARF e Aux/IAA (TIWARI et al., 2001).

Auxina ou ácido indol acético (IAA) regula diversas respostas celulares e de desenvolvimento em plantas, incluindo divisão, expansão e diferenciação celular, este fitormônio age na expressão de muitos genes cujos produtos provavelmente participantes das respostas de desenvolvimento (ABEL e THEOLOGIS, 1996).

BÁSTIAN et al. (1998) caracterizaram por cromatografia gasosa-espectrometria de massa (GC-MS) os fitormônios ácido indol acético e giberelinas liberados em meio de cultura pelos diazotrofos endofíticos *G. diazotrophicus* e *H.seropedicae*. A produção de pequenas quantidades deste composto pelo endófito sugere uma explicação para a maior frequência de genes relacionados a resposta a auxina na biblioteca de plantas inoculadas com *H.seropedicae*.

TABELA 6 – GENES DE RESPOSTA A FITORMÔNIOS DA BIBLIOTECA RRCH5SN

Clone	Id	Comprimento da sequência (bp)	Número de bases Alinhadas (bp)	Descrição	Cromosomo	Número de acesso	GO_id
RRCH5SN01-D01.b1	99%	234	230/231	Fator transcricional relacionado a patogênese e proteína contendo domínio ERF (provável proteína ligadora a elemento responsivo a etileno) -proteína contendo domínio AP2	3	<u>016-063-E02 </u> Os03g0183300 LOC_Os03g08500	<u>GO:0003677</u> Ligação de DNA
RRCH5SN07-F09	98%	232	225/229	Metionina sintase responsiva a etileno (fragmento) similar a Q9SWV3 (provável 5-metiltetrahidropterilroiltriglutamato-homocisteína metiltransferase) similar a proteínas pertencentes à família metionina sintase independente de vitamina B12	12	GENE ID: 4352833 Os12g0624000	GO:0003871: atividade 5-metiltetrahidropterilroiltriglutamato-homocisteína S-metiltransferase
RRCH5SN09-C06	98%	351	345/349	Proteína responsiva a auxina IAA7 (proteína 7 induzida por ácido indol acético) similar a O24408	2	<u>GENE ID:</u> <u>4328796</u> <u>Os02g0228900</u> NP_001046357.1	Não encontrada
RRCH5SN09-D03	100%	409	406/406	Elemento ligante a fator 5 responsivo a etileno (ERF5) proteína tipo EREBP Similar a O80341 -fatores relacionados a AP2/EREBP	7	<u>GENE ID:</u> <u>4343912</u> <u>Os07g0617000</u> NP_001060280.1	GO:0003700: atividade de fator de transcrição

Continuação tabela 6

Clone	Id	Comprimento da sequência (bp)	Número de bases Alinhadas (bp)	Descrição	Cromosomo	Número de acesso	GO_id
RRCH5SN09-H06.b2	98%	399	292/296	Factor 3 responsivo a auxina (ARF 2)	1	<u>AK103776 </u> Os01g0670800 <u>LOC_Os01g48060.1</u>	GO:0003700: atividade de fator de transcrição
RRCH5SN14-F05	100%	122	90/90	Metionina sintase responsiva a etileno (fragmento) similar a Q9SWV3	12	<u>GENE ID: 43 52833</u> <u>Os12g0624000</u> NP_001067314.1	GO:0003871: atividade metiltetrahydropteroilglutamato # homocisteína metiltransferase
RRCH5SN24-E02.b1	97%	101	96/98	Gene membro da família Aux/IAA OslAA19 – responsivo a auxina	5	<u>014-099-C02 </u> Os05g0559400 <u>LOC_Os05g48590.1</u>	GO:0005634: núcleo

Id = identidade com o banco de dados

Go_Id: número de acesso GO do banco de dados *Gene Ontology*

TABELA 7 – GENES DE RESPOSTA A FITORMÔNIOS DA BIBLIOTECA RRCH5SN

Clone	Id	Comprimento da sequência (bp)	Número de bases Alinhadas (bp)	Descrição	Cromosomo	Número de acesso	GO_id
RRSH0SN05-A01.b3	100%	505	502/502	Proteína da família AUX/IAA	1	<u>GENE ID:</u> <u>4326957</u> <u>Os01g0927600</u> <u>NP_001045270.1</u>	GO:0003677: ligação de DNA
RRSHSN14-F05	98%	405	368/372	Metionina sintase responsiva a etileno (fragmento)	12	<u>gb EF576403.1 </u> <u>Os12g0624000</u> <u>NP_001067314.1</u>	GO:0003871: atividade 5-metiltetraidrop teroiltriglutamato homocisteína-S-metiltransferase

Id = identidade com o banco de dados

Go_Id: número de acesso GO do banco de dados *Gene Ontology*

5.3.3 Genes mais freqüentes encontrados em ambas bibliotecas

Genes codificando metalotioneínas foram os mais representados tanto na biblioteca controle quanto na biblioteca inoculada com *H.seropedicae* por 5 dias (tabela 8 e 9).

Metalotioneínas (MTs) constituem uma superfamília evolutivamente conservada, de baixa massa molecular, são proteínas ricas em cisteínas que podem ligar a metais por meio de grupos tióis de seus resíduos de cisteína (COYLE et al., 2002). MTs são amplamente encontradas em diversos organismos (COYLE et al., 2002; COBBETT e GOLDSBROUGH, 2002). Estudos em plantas demonstraram que MT além de estarem envolvidas na manutenção da homeostase e detoxificação de metais (COBBETT e GOLDSBROUGH, 2002), também estão envolvidas na captura de espécies reativas de oxigênio (ROS) (AKASHI et al., 2004; WONG et al., 2004).

Baseado em análises estruturais e filogenéticas, a família de OsMT foi classificada em classe I e classe II. A classe I, por sua vez, foi subdividida em 4 tipos, esta categorização das classes é baseada no arranjo dos resíduos de cisteína em sua seqüência de aminoácidos (GUO et al., 2003). A maioria das metalotioneínas encontradas tanto em dicotiledôneas como em monocotiledôneas pertencem à classe I (ZHOU et al., 2006).

Na biblioteca inoculada com *H.seropedicae* foram encontradas 11 ESTs de metalotioneínas: uma tipo 2B, uma tipo 3A e 9 metalotioneínas do tipo 1A. Na biblioteca controle, foram encontradas 11 ESTs de metalotioneínas do tipo 1A e duas do tipo 4A.

O gene Os01g0974200, encontrado na biblioteca inoculada, codifica uma metalotioneína tipo 2B, denominada RicMT ou OsMT-I-2b e pertencente à classe I. Este tipo de metalotioneína é altamente expressa em tecidos de arroz em divisão celular e sofrendo alongação celular (YU et al., 1998). Possuindo à montante do gene, elementos regulatórios, com elementos responsivos a metais (MREs) (ZHOU et al., 2006).

Na biblioteca inoculada, foi encontrado um gene Os01g0200700 que codifica uma metalotioneína tipo 3A, também pertencente à classe 1, denominada OsMT-I-3a ou MT3A. Esta metalotioneína foi encontrada induzida em folhas de arroz que receberam amônio como fonte de nitrogênio (ZHU et al., 2006).

Na biblioteca controle foi encontrado um gene Os12g0570700 em um contig, contendo 2 ESTs codificando OsMT-I-4A.

Zhou et al. (2006) analisou a expressão de metalotioneínas de arroz pelo método de “Northern blotting” e mostrou alto nível de expressão de OsMT-1-3a, OsMT-1-4a e OsMT-1-4b, em raízes, mostrando que existem padrões de expressão de cada OsMT, característicos dos tecidos ou fases de desenvolvimento. A expressão diferencial destes genes sugere que metalotioneínas são importantes para evitar o dano oxidativo pelo sequestro de íons metálicos nas raízes (COBBETT e GOLDSBROUGH, 2002).

Dois contigs, cada um com 9 ESTs, resultante do alinhamento de ESTs de das bibliotecas RRCH5SN e RRSN, representam o gene Os11g0704500, que codifica uma metalotioneína do tipo 1A (classe I) denominada OsMT-I-1a ou MT1A. Na biblioteca RRSN o gene Os12g0568500, que codifica uma proteína similar a esta, foi identificado em um contig com duas ESTs. JIN et al. (2006) encontraram alta expressão de OsMT-I-1a quando tecidos de arroz foram expostos a estresse abiótico como seca, alta concentração salina (NaCl, NaHCO₃) e metais (CuCl₂, ZnCl₂, CdCl₂), tanto em folhas quanto em raízes de arroz (JIN et al., 2006). HSIEH et al. (1995) mostrou que o gene que codifica para OsMT-I-1A é induzido por estresse após tratar tecidos de arroz também com metais pesados e exposição a outros estresses abióticos como choque por calor e limitação de fonte de sacarose, observando também aumento em sua expressão.

Análises de expressão de ESTs de arroz durante a interação com fungo *Magnaporthe grisea* revelaram uma expressão aumentada em duas vezes de metalotioneína do tipo 1 em resposta a patógenos, porém também foi observada expressão constitutiva desta proteína mesmo em folhas saudáveis de arroz não inoculadas (KIM, ANN e LEE, 2001)

O nível de expressão de ESTs encontradas acima será determinado futuramente e a função destes genes será estudada em relação à interação planta-bactéria.

TABELA 8 – METALOTIONEÍNAS ENCONTRADAS NA BIBLIOTECA RRCH5SN

Clone	Id	Comprimento da sequência (bp)	Número de bases Alinhadas (bp)	Descrição	Cromosomo	Número de acesso	GO_id
RRCH5SN02-F8	99%	139	135/136	RicMT (proteína tipo metalotioneína 2B) classe I idêntica a O04107	1	<u>GENE ID:</u> <u>4326409</u> <u>Os01g0974200</u> <u>NP_001045552.1</u>	GO:0046872: ligação a íon metal
RRCH5SN10-A02.b1	92%	131	102/110	Proteína tipo Metalotioneína-3A classe I	1	<u>J100079J19 </u> <u>Os01g0200700</u> <u>A1YTM8</u>	Não encontrada
RRCH5SN02-C08 RRCH5SN02-D04 RRCH5SN03-E08 RRCH5SN07-A07 RRCH5SN09-C11 RRCH5SN09-H02 RRCH5SN14-B02 RRCH5SN21-E03 RRCH5SN21-H03	97%	535	528/543	Proteína tipo Metalotioneína-tipo 1 Proteína tipo Metalotioneína 1A classe1	11	<u>dbj AK062653.1 </u> <u>Os11g0704500</u> <u>NP_001068544.1</u>	GO:0046872: ligação de íon metal

Id = identidade com o banco de dados

Go_Id: número de acesso GO do banco de dados *Gene Ontology*

TABELA 9 – METALOTIONEÍNAS ENCONTRADAS NA BIBLIOTECA RRSN

Clone	Id	Comprimento da sequência (bp)	Número de bases Alinhadas (bp)	Descrição	Cromosomo	Número de acesso	GO_id
RRSHSN01-D07 RRSHSN21-C05	97%	319	302/311	Proteína tipo metalotioneína tipo 1	12	emb CT831673.1 ↓ Os12g0568500 NP_001067066.1	GO:0046872: ligação de íon metal
RRSHSN15-H05 RRSHSN29-F04	100%	344	337/337	Proteína tipo metalotioneína	12	GENE ID: 4352580 Os12g0570700 NP_001067081.1	GO:0046872: ligação de íon metal
RRSHSN07-G02 RRSHSN14-B02 RRSHSN14-G06 RRSHSN15-A05 RRSHSN22-F11 RRSHSN24-C06 RRSHSN24-C11 RRSHSN24-G08 RRSHSN26-E07	97%	488	477/491	Proteína tipo metalotioneína tipo 1 Idêntica a Q40633	11	dbj AK120630.1 Os11g0704500 NP_001068544.1	GO:0046872: ligação de íon metal

Id = identidade com o banco de dados

Go_Id: número de acesso GO do banco de dados *Gene Ontology*

5.3.4 Prováveis novos genes encontrados neste estudo

Algumas ESTs quando analisadas através do BLASTN foram caracterizadas inicialmente como DNA genômico de arroz. Entretanto utilizando o programa BLASTX e o banco de dados nr (www.ncbi.nlm.nih.gov), foi possível concluir que algumas destas seqüências provavelmente codificam proteínas similares a outras encontradas em diversos organismos. Este resultado sugere que se tratam de novos genes de arroz. Foi encontrado um total de 5 genes novos em ambas bibliotecas (tabela 10).

Na biblioteca RRSHSN foi encontrado um contig, agrupando 3 ESTs, que codifica para uma provável citocromo P450 monoxigenase encontrada em milho. Nesta mesma biblioteca foi encontrada uma EST correspondente a uma provável proteína como dedo de zinco de *Hyacinthus orientalis* e uma correspondente a uma provável poliubiquitina de *Adonis aestivalis*.

Na biblioteca inoculada foram encontradas 2 ESTs agrupadas em um contig que inicialmente identificado como cDNA de arroz não caracterizado. Posteriormente, a análise do BLASTX revelou que se trata de gene codificando uma provável proteína expressa em *Vitis vinifera*. Também foi encontrada uma seqüência de DNA genômico de arroz que posteriormente foi identificada como um possível RNA mensageiro de milho, ainda não caracterizado.

Estas seqüências que transcrevem possíveis novos genes terão suas funções futuramente caracterizadas e seu papel em relação à interação planta-bactéria poderá ser determinado.

TABELA 10 – PROVÁVEIS NOVOS GENES DE *ORYZA SATIVA* ENCONTRADOS NESTE TRABALHO

Clone	Id	Comprimento da sequência (bp)	Número de bases Alinhadas (bp)	Descrição	Cromosomo	Número de acesso	GO_id
RRSHSN14-C08 RRSHSN20-B10 RRSHSN21-A08	99%	323	318/320	(cDNA) provável citocromo P450 monoxigenase de milho (fragmento)	9	dbj AK119873.1 pir T02955	Não encontrada
RRSH0SN07-C05.b1	94%	198	188/198	(DNA genômico) Provável proteína "zinc finger" [<i>Hyacinthus orientalis</i>]	10	gb AC145127.1 gb AAQ89709.1 	Não encontrada
RRSH0SN14-D10.b2	99%	426	409/412	(DNA genômico) Provável poliubiquitina [<i>Adonis aestivalis</i> var. <i>palaestina</i>]	5	gb AC152972.1 gb ABY60454.1 	Não encontrada
RRCH5SN02-E10 RRCH5SN08-C02	99%	404	400/401	(cDNA) proteína hipotética [<i>Vitis vinifera</i>]	9	dbj AK289006.1 emb CAN61119.1 	Não encontrada
RRCH5SN05-A09	100%	471	381/381	(DNA genômico) [mRNA <i>Zea mays</i> 409/475 (86%)]	2	dbj AP008208.1 	Não encontrada

■ Análise por BLASTX contra banco nr (www.ncbi.nlm.nih.gov)

Id = identidade com o banco de dados

Go_Id: número de acesso GO do banco de dados *Gene Ontology*

5.4 CONSIDERAÇÕES FINAIS

Ainda não está claro quais mecanismos estão envolvidos no estabelecimento de interações planta-bactéria, e quais moléculas participam da sinalização entre a planta e a bactéria. A planta provavelmente controla a colonização da bactéria pelo envio de sinais moleculares e/ou provendo melhores condições fisiológicas para a sobrevivência da bactéria (NOGUEIRA et al., 2001).

A investigação da função biológica de ESTs expressas preferencialmente na biblioteca de cDNA inoculada com *H.seropedicae* pode conduzir à identificação de novos genes envolvidos nesta associação.

Os resultados deste trabalho apontam prováveis genes de arroz que podem estar envolvidos com esta interação. Trabalho futuro envolvendo determinação do nível de expressão destes genes em plantas inoculadas e não inoculadas e produção de plantas com estes genes encontrados deverão esclarecer suas funções.

Uma série de genes de cana-de-açúcar envolvidos com o desenvolvimento da interação planta-bactéria já foi caracterizada (NOGUEIRA et al., 2001; LAMBAIS et al., 2001). Os produtos de alguns destes genes operam em vias de sinalização de etileno e auxina (CAVALCANTE et al., 2007; WANG, et al., 2007; CAO et al., 2006; GUO et al., 2004; BÁSTIAN, et al., 1998). Na biblioteca de cDNA de raízes de arroz inoculada com *H.seropedicae* foram identificados em número substancialmente maior de genes destas vias, sugerindo sua participação no desenvolvimento da interação entre *H. seropedicae* e plântulas de arroz.

6. CONCLUSÕES

- Foram seqüenciadas as extremidades 5' de 4608 clones de duas bibliotecas de cDNA de arroz: uma inoculada por 5 dias com *H.seropedicae* (RRCH5SN) e outra sem tratamento(RRSHSN);
- 2314 seqüências possuíam mais de 100 pares de base com Phred maior que 15;
- As seqüências da biblioteca RRCH5SN continha 701 singlets e 54 contigs identificados, totalizando 755 ESTs únicas;
- Na biblioteca não inoculada, RRSHSN, foram identificados 560 singlets e 62 contigs, totalizando 622 ESTs únicas;
- Das 755 ESTs únicas da biblioteca RRCH5SN, 82% das seqüências possuem identidade igual ou superior a 90% com genes de arroz;
- 72% das 622 ESTs únicas da biblioteca RRSHSN possuem identidade igual ou superior a 90% com genes de arroz;
- Foram identificados genes expressos exclusivamente na biblioteca de cDNA de raízes de arroz inoculadas com *H.seropedicae*. Estes são: dois transportadores de nitrato, dois codificando proteínas contendo o domínio WD40 e um gene codificando proteína SEP2 de resposta a estresse;
- Foram encontrados genes de resposta aos hormônios vegetais etileno e auxina sendo expressos preferencialmente na biblioteca RRCH5SN;
- Genes codificando metalotioneínas da classe I foram os mais freqüentemente encontrados em ambas as bibliotecas;
- Foram encontrados 5 prováveis novos genes de arroz.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABEL, S.; THEOLOGIS, A. Early genes and auxin action. **Plant Physiol**, v.111, p.9-17, 1996.
- ADAMS, M.; KERLAVAGE, A.R.; FLEISCHMANN, R.D; FULDNER, R.A.; BULT, C.J.; LEE, N.H.; KIRKNESS, E.F.; WEINSTOCK, K.G.; GOCAYNE, J.D., WHITE, O., et al. Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. **Nature**. v.377, p.173–174, 1995.
- AKASHI, K., NISHIMURA, N., ISHIDA, Y., and YOKOTA, A. Potent hydroxyl radical-scavenging activity of drought-induced type 2 metallothionein in wild watermelon. **Biochem. Biophys. Res. Com.** v.323, p. 72-78, 2004.
- ALTSCHUL, S.F.; MADDEN, T.L.; SCHÄFFER, A.A.; et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v.25, n.17, p.3389-3402, 1997.
- BALDANI, V.L.D.; BALDANI, J.I.; DÖBEREINER, J. Inoculation of rice plants with the endophytic diazotrophs *Herbaspirillum seropedicae* and *Burkholderia spp.* **Biology and Fertility of Soils**, v. 30, p. 485–491, 2000.
- BALDANI, V.L.D. Efeito da inoculação de *Herbaspirillum spp.* no processo de colonização e infecção de plantas de arroz e ocorrência e caracterização parcial de uma nova bactéria diazotrófica seropédica. . **Tese de doutorado** UFRRJ, p. 234, 1996.
- BALDANI, J.I.; BALDANI, V.L.D.; SELDIN, L; DOBEREINER, J. Characterization of *Herbaspirillum seropedicae* gen. nov., sp. nov., a root-associated nitrogen fixing bacterium. **Intl. J. Sys. Bact.**, v. 36, p. 86-93, 1986.
- BARRAQUIO, W. L., M. R. DE GUZMAN, M. BARRION, E I. WATANABE. Population of aerobic heterotrophic nitrogen-fixing bacteria associated with wetland and dryland rice. Appl. **Environ. Microbiol.** v.42, p.124-128, 1982.
- BÁSTIAN, F., COHEN, A., PICCOLI, P., et al. production of indole-3-acetic acid and gibberelins A1 e A3 by *Acetobacter diazotrophicus* and *Herbaspirillum seropedicae* in chemically-defined culture media. **Plant Growth Regul.** v.24, p.7-11, 1998.

BENHAMOU, N.; KLOEPPER J.W.; QUADT-HALLMAN, A.; TUZUN, S.. Induction of defence-related ultrastructural modifications in pea root tissues inoculated with endophytic bacteria. **Plant Physiol.** v.112, p. 919–929, 1996.

BLADERGROEN, M. R.; H. P. SPAINK. Genes and signal molecules involved in the rhizobia-Leguminosae symbiosis. **Curr. Opin. Plant Biol.** v.1, p.353-359, 1998.

BOCKMAN, O.C.. Fertilizers and biological nitrogen fixation as sources of plant nutrients: perspectives for future agriculture. **Plant and Soil**, v.194, p. 11-14, 1997

BOCKMAN, O.C.; KAARSTAD, O.; LIE, O.H.; RICHARDS, I. Agriculture and fertilizers: fertilizer in perspective. **Norsk Hydro, Oslo**.1990.

BODDEY, R.M.; OLIVEIRA. O.C. de; URQUIAGA, S.; et al. Biological nitrogen fixation associated with sugar cane and rice: contributions and prospects for improvement. **Plant and Soil**, v. 174, p. 195-209, 1995.

BRENCIC, A.; WINANS, S.C. Detection of and Response to Signals Involved in Host-Microbe Interactions by Plant-Associated Bacteria. **Microbiol. Molec. Biol. Rev.** v.69, p. 155-194, 2005.

BROUGHTON, W. J.; S. JABBOURI; X. PERRET. Keys to symbiotic harmony. **J. Bacteriol.** v.182, p.5641-5652, 2000.

BROUN, P. Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in *Arabidopsis*. **Curr. Opin. Plant. Biol.** v.8, p.272-279, 2005.

BRUSAMARELO, L.C. **Sequenciamento e análise de seqüências expressas de arroz (ESTs) de arroz (*Oryza sativa*) inoculado com *Herbaspirillum seropedicae***. 69 f. Dissertação (Mestrado em Genética) – Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2007.

BULL, C. T.; WELLER, D. M. E THOMASHOW, L. S. Relationship between root colonization and suppression of *Gaeumannomyces graminis* var. *tritici* by *Pseudomonas fluorescens* strain 2-79. **Phytopathol.** v.81, p. 954-959, 1991.

BURRIS, R. H. Nitrogenases. **J. Biol. Chem.**, v. 266, p. 9339-9342, 1991.

CANUTO, E.L.; OLIVEIRA, A.L.M.; REIS, V.M.; et al. Evaluation of the biological nitrogen fixation contribution in sugarcane plants originated from seeds and inoculated with nitrogen-fixing endophytes. **Brazilian Journal of Microbiology**, v. 34, p. 62-64, 2003.

CAO, Y.; SONG, F.; GOODMAN, R.M.; ZHENG, Z. Molecular characterization of four transcriptional factors and their expression in response to biotic and abiotic stress. **Jour. Plant. Physiol.** v. 163, p.1167-1178, 2006

CAVALCANTE, J.J.V.; VARGAS, C.; NOGUIERA, E.M., VINAGRE, F.; SCHWARCZ, K.; BALDANI, J.I.; FERREIRA, P.C.G.; HEMERLY, A.S. Members of the ethylene signalling pathway are regulated in sugarcane during the association with nitrogen-fixing endophytic bacteria. **Jour. Experim. Botany.** v. 58, p. 673-686, 2007.

CHOUDHURY, A.T.M.A., KENNEDY, I.R. Prospects and potentials for systems of biological nitrogen fixation in sustainable rice production. **Biol. Fertil. Soils.** v.39, p. 219-227, 2004.

COBBETT, C.; GOLDSBROUGH, P. Phytochelatins and metallotioneins: roles in heavy metal detoxification and homeostasis. **Annu. Rev. Plant. Biol.** v. 53, p. 159-182, 2002.

COYLE, P.; PHILCOX, J.C.; CAREY, L.C.; ROFE, A.M.; Metallotionein: the multipurpose protein. **Cell. Mol. Life Sci.** v. 59, p. 627-647, 2002.

CRUZ, L.M.; SOUZA, E.M.; WEBER, O.B. et al. 16S ribosomal DNA characterization of nitrogen-fixing bacteria isolated from banana (*Musa* spp.) and pineapple (*Ananas comosus* (L.) Merrill). **Applied and Environmental Microbiology**, v. 67, n.5, p. 2375-2379, 2001.

CUI, Y.; XINYE, Z.; ZHOU, Y.; LIN, T.; ZHANG, L.; ZHUGE, Q.; CAI, Y.; HUANG, M. Identification and expression analysis of EST-based genes in the bud of *Lycoris longituba*. **Geno. Prot. Bioinfo.** v.2, p.43-46, 2004.

DIXON, R.; KAHN, D. Genetic regulation of biological nitrogen fixation. **Nature reviews Microbiology.** v.2, p.621-631, 2004.

DOBEREINER, J.; BALDANI, V.L.D.; REIS, V.M. Endophytic occurrence of diazotrophic bacteria in non-leguminous crops. **In Azospirillum VI and Related**

Microorganisms. Eds. Fendrik, I.; del Gallo, M.; VANDERLEYDEN, J.; de ZAMAROCZY, M. p. 3-14. Springer, Berlin, Heidelberg, Germany, 1995.

EWING, B.; GREEN, P. Base-Calling of Automated Sequencer Traces Using *Phred*. II. Error Probabilities. **Genome Research**, v.8, n.3, p.186-194, 1998.

EWING, B.; HILLIER, L.; WENDL, M.C. et al. Base-calling of automated sequencer traces using *Phred*. I. Accuracy Assessment. **Genome Research**, v.8, n.3, p.175-185, 1998.

EULGEM, T.; RUSHTON, P.J., SHMELZER, E.; HAHLBROCK, K.; SOMSSICH, I.E. Early nuclear events in plant defence signaling: rapid activation by WRKY transcription factors. **EMBO J.** v.18, p.4689-4699, 1999.

FINDANTISEF, A.T.; STOUT, M.L.; THALER, J.S.; DUFFEY, S.S. Signal interaction in pathogen and insect attack: expression of lipoxygenase, proteinase inhibitor II, and pathogenesis related protein P4 in tomato, *Lycopersicon esculentum*. **Physiol Mol. Plant Pathol.** v.54, p. 97-114, 1999.

FUENTES-RAMIREZ, L.E.; JIMENEZ-SALGADO, T.; ABARCA-OCAMPO, I.R.; CABALLERO-MELLADO, J. *Acetobacter diazotrophus*, an indolacetic acid-producing bacterium isolated from sugarcane cultivar in Mexico. **Plant Soil.** v.154, p. 145-150, 1993.

FUJIMOTO, S.Y.; OHTA, M; USUI, A.; SHINSHI, H.; OHME-TAKAGI, M. Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC Box-mediated gene expression. **The Plant Cell**, v.12, p.393-404, 2000.

FUJIWARA, S.; TANAKA, N.; KANEDA, T.; TAKAYAMA, S.; ISOGAI, A.; CHE, F. Rice DNA microarray-based gene expression profiling of the response to flagellin perception in cultured rice cells. **Mol. Plant Microbe Interact.** v.17, p. 986-998, 2004.

GALÁN, J.E.; WOLF-WATZ, H. Protein delivery into eukaryotic cells by type III secretion machines. **Nature.** v.144, p. 567-573, 2006.

GOFF, S. A.; RICKE, D.; LAN, T.; et al. A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. japonica). **Science**, v. 296, p. 92-100, 2002.

GUO, H.; ECKER, J.R. The ethylene signaling pathway: new insights. **Curr. Opin. Plant. Biol.** v.7, p.40-49, 2004.

GYANESHWAR, P, JAMES, E. K, REDDY P. M., E LADHA, J. K. *Herbaspirillum* colonization increases growth and nitrogen accumulation in aluminium-tolerant rice varieties. **New Phytologist** . v.154 , n.1, p.131-145, 2002.

HAGEN, G.; GUILFOYLE, T. Auxin-responsive gene expression: genes, promoters and regulatory factors. **Plant Mol. Biol.** v.49, p.373-385, 2002.

HEDDAD, M.; ADAMSKA, I. Light stress-regulated two helix proteins in *Arabidopsis thaliana* related to the chlorophyll a/b-binding gene family. **PNAS**, v.97, p.3741-3746, 2000.

HSIEH, H.; LIU, W; HUANG, P.C. A novel stress-inducible metallothionein-like gene from rice. **Plant Molecular Biology**. v.28, p.381-389, 1995.

HOUDE, M; BELCAID, M.; OUELLET, F.; DANYLUK, J.; MONROY, A.F.; DRYANOVA, A.; GULICK,P.; BERGERON, A; LAROCHE, A; LINKS, M.G.; MACCARTHY, K.; CROSBY, W.L.; SARHAN, F. Wheat EST resources for functional genomics of abiotic stress. **BMC Genomics**, v. 7, p.1-22, 2006.

HOWARD, J.B.; REES, D.C. Structural basis of biological nitrogen fixation. **Chem Rev** v.96, p.2965-2985, 1996.

HUANG, X.; MADAN, A. CAP3: A DNA Sequence Assembly Program. **Genome Research**, v 9, p.868-877, 1999.

IRRI. **Rice today**, v.5, n.3. Metro Manila, Filipines, 2006.

ITOH, T; TANAKA, T; BARRERO, R. A. YAMASALI, C.; FUJI, Y.; HILTON, P.B.; ANTONIO, B.A.; AONO, H.; APWEILER, R. et al., Curate genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*. **Genome Res.** v.17, p. 175-183, 2007.

JAMES, E. K.; GYANESWAR, P.; MATHAN, N.; BARRAQUIO, W. L.; REDDY, P. M.; IANNETTA, P. P. M.; OLIVARES, F. L. E LADHA, J. K. Infection and colonization of rice seedlings by the plant growth-promoting bacterium *Herbaspirillum seropedicae* Z67. **Mol. Plant-Microbe Interact.** v.15, p.894-906, 2002.

JAMES, E. K., E F. L. OLIVARES. Infection and colonization of sugarcane and other graminaceous plants by endophytic diazotrophs. **Crit. Rev. Plant Sci.** v.17, p.77-119, 1998.

JAMES, E.K.; OLIVARES, F.L.; BALDANI, J.I. E DÖBEREINER, J. *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue in leaves of *Sorghum bicolor* L. Moench. **J. Exp. Bot.** v.48, p.785-797, 1997.

JANTASURIYARAT, C., GOWDA, M., HALLER, K., et al. Large-scale identification of expressed sequence tags involved in rice and blast fungus interaction. **Plant Physiology** . v.138, p. 105-115, 2005.

JIN, S.; CHENG, Y.; GUAN, Q.; LIU, D.; TAKANO, T.; SHENKUI, S. A metallotionein-like protein of rice (rg-MT) functions in E.coli and its gene expression is induced by abiotic. *Biotechnol Let.* v.28, p. 1749-1753, 2006.

KARGER, A.E.; HARRIS, J.M; GESTELANG, R.F. Multiwavelength fluorescent detection for DNA sequencing using capillary electrophoresis. **Nucleic Acid Res.**, v.19, p.4955-4962, 1991.

KENNEDY I.R., TCHAN Y.T. Biological nitrogen fixation in non-leguminous field crops: Recent advances. **Plant Soil**; v.141, p.93-118, 1992.

KIKUCHI, S. SATOH, K., NAGATA, T., et al. (The rice full-length cDNA consortium) Collection, mapping and annotation of over 28,000 cDNA clones from *japonica* rice. **Science** v.301, p.376-378, 2003.

KIM, S.; AHN, P; LEE, Y. Analysis of genes expressed during rice-*Magnaporthe grisea* interactions. **Molecular Plant-Microbe Interactions**. v.14, p. 1340-1346, 2001

KIM, C.Y.; LEE, S.; PARK, H.C.; BAE, C.G.; CHEONG, Y.H; CHOI, Y.C.; HAN, C.; LEE, S.Y.; LIM, C.O.; CHO, M.J. Identification of rice blast fungal elicitor-responsive genes by differential display analysis. **Molecular Plant-Microbe Interactions**. v.13, p.470-474, 2000.

LADHA J. K., P. M. REDDY. Nitrogen fixation in rice systems: state of knowledge and future prospects, **Plant and Soil** v.252, n.1, p.151-167, 2003.

LADHA, J. K., E P. M. REDDY. Extension of nitrogen fixation to rice necessity and possibilities. **GeoJournal** v.35, p.363-372, 1995.

LAMBAIS, M.R. *In silico* differential display of defense-related expressed sequence tags from sugarcane tissues infected with diazotrophic endophytes. **Genetics and Molecular Biology**, v.24, p.103-111, 2001.

LA ROTA, M.; SORRELLS, M. E. comparative DNA sequence analysis of mapped wheat Ests reveals the complexity of genome relationships between rice and wheat. **Funct. Integr. Genomics**. v.4, p.34-46, 2004.

LAI, J.; DEY, N.; KIM, C.; BHARTI, A.K.; RUDD, S.; MAYER, K.F.X.; LARKINS, B.A.; BECRAFT, P.; MESSING, J. Characterization of the maize endosperm transcriptome and its comparison to the rice genome. **Genome Research**.v.14, p. 1932-1937, 2004.

LISCUM, E.; REED, J.W. Genetics of Aux/IAA and ARF action in plant growth and development. **Plant. Molecular Biology**. V. 49, p. 387-400, 2007.

MALIK, K. A., B. RAKHSHANDA, S. MEHNAZ, G. RASUL, M. S. MIRZA, E S. ALI. Association of nitrogen-fixing, plant growth promoting rhizobacteria (PGPR) with Kallar grass and rice. **Plant Soil** . v.194, p.37-44, 1997.

MATSUMOTO, T.; WU, J.; KANAMORI, H. et al (International Rice Genome Sequencing Project). The map-based sequence of the rice genome. **Nature**, v.436, p.793-800, 2005.

MCDOWELL, J.M.; DANGL, J.L. Signal transduction in the plant immune response. **Trends in Biochemical Science**, v.25, n.2, p. 79-82, 2000.

MERRICK, M. J. Regulation of nitrogen fixation genes in free-living and symbiotic bacteria. In: STACEY, G.; BURRIS, R. H.; EVANS, H. J. (Eds.). **Biological nitrogen fixation**, 1992.

MICHALEK, W.; WESCHKE, W.; PLEISSNER, K.P.; GRANER, A. EST analysis in barley defines a unigene set comprising 4,000 genes. **Theor Appl Genet** .v.104, p.97–103, 2002.

MILLER, J.F.; MEKALANOS, J.J.; FALKOW, S. Coordinate regulation and sensory transduction in the control of bacterial virulence. **Science**, Washington, v. 243, p. 916-922, 1989.

MUTHUKUMARASAMY, R.; REVATHI, G.; LAKSHMINARASIMHAN, C. Diazotrophic associations in sugar cane cultivation in South India. **Tropical Agriculture**, v.76, p.171–178, 1999.

NOGUEIRA, E.M., VINAGRE, F., MASUDA, H.P., VARGAS, C. et al. Expression of sugarcane genes induced by inoculation with *Acetobacter diazotrophicus* and *Herbaspirillum rubrisubalbicans*. **Gen. and Mol. Biol.** v.24, n.1-4, p.199-206, 2001.

OGIHARA, Y.; MOCHIDA, K.; NEMOTO, Y.; MURAI, K.; YAMAZAKI, Y., SHIN-IT; KOHARA, Y. . Correlated clustering and virtual display of gene expression patterns in the wheat life cycle by large-scale statistical analyses of expressed sequence tags. **Plant J** v.33, p.1001–1011, 2003

OHLROGGE, J.; BENNING, C. Unraveling plant metabolism by EST analysis. **Curr. Opin. Plant Biol.** v.3, p.224-228, 2000

OHME-TAKAGI, M.; SUZUKI, K.; SHINSHI, H. Regulation of ethylene-induced transcription of defense genes. **Plant Cell Physiol.** v. 41, p. 1187-1192, 2000.

OHME-TAKAGI, M.; SHINSHI, H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. **The Plant Cell.** V. 7, p. 173-182, 1995.

OHTA, H.; SHIRANO, Y.; TANAKA, K.; MORITA, Y.; SHIBATA, D. CDNA cloning of rice lipoxygenase L-2 and characterization using an active enzyme expressed from the cDNA in *Escherichia coli*. **Eur. J. biochem.** v.206, p. 331-336, 1992.

OHTA, H.; SHIDA; K. PENG, Y.; FURUSAWA, I.; SHISHIYAMA, J.; AIBARA, S.; MORITA, Y. A lipoxygenase pathway is activated in rice after infection with rice blast fungus *Magnaporthe grisea*. **Plant physiol.** v. 97, p. 94-98, 1991.

OHTA, H.; SHIDA; K. PENG, Y.; FURUSAWA, I.; SHISHIYAMA, J.; AIBARA, S.; MORITA, Y. The occurrence of lipid hydroperoxide-decomposing activities in rice and relationship of such activities to the formation of antifungal substances. **Plant Cell Physiol.** v.31, p.1117-1122, 1990.

OLIVARES, F.L.; BALDANI, V.L.D.; REIS, V.M. BALDANI, J.I.; DÖBEREINER, J. Occurrence of endophytic diazotrophs *Herbaspirillum* spp. in roots, stems and leaves predominantly of Gramineae. **Biol. Fertil. Soil.** v.21, p. 197-200, 1996.

OHYANAGI, et al. The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* spp. *japonica* genome information. **Nucleic Acid Research.** v.34, p.741-744, 2006.

PENMETSA, R.V.; COOK, D.R. A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. **Science.** v. 275, p.527-530, 1997.

PERRET, X., C.; STAEHELIN; W. J. BROUGHTON. Molecular basis of symbiotic promiscuity. **Microbiol. Mol. Biol. Rev.** v.64, p.180-201, 2000.

PIMENTEL, J. P.; OLIVAR ES, F.; PITARD, R. M.; URQUIAGA, S.; AKIBA, F.; DOBEREINER, J. Dinitrogen fixation and infection of grass leaves by *Pseudomonas rubrisubalbicans* and *Herbaspirillum seropedicae* **Plant. Soil.; Dordrecht**, v. 137, p. 61-65, 1991.

POSTGATE, J. R. The fundamentals of nitrogen fixation. Cambridge, Cambridge Univ.Press, 1982.

PÜHLER, A ;ARLAT, M.; BECKER, A; GÖTTFERT, M.; MORRISSEY, J.P.; O'GARA, F. What can bacterial genome research teach us about bacteria-plant interactions? **Curr Opin Plant Biol**, v. 7, p. 137-147, 2004.

QUISPEL, A. A search of signals in endophytic microorganisms. In: **Molecular signals in Plant-Microbe Communications. (D.P.S. Verma, Ed.) p.471-491. CRC Press, Boca Raton, FL, 1992.**

REDDY, A.R.; RAMAKRISHNA, W.; SEKHAR, A.C. et al. Novel genes are enriched in normalized cDNA libraries from drought-stressed seedlings of rice (*Oryza sativa* L. subsp. *indica* cv. Nagina 22). **Genome**, v.45, p.204-211, 2002.

RONCATO-MACCARI, L. D. B., RAMOS, H. J. O., FO, P., ALQUINI, YEDO, CHUBATSU, L. S., YATES, M G, RIGO, L U, STEFFENS, M. B. R., SOUZA, E. M. Endophytic *Herbaspirillum seropedicae* expresses *nif* gene in gramineous plants. **FEMS Microbiology Ecology**, v.45, p.39 – 47, 2003a.

RONCATO-MACCARI, L.D.B, RAMOS, H.J.O, PEDROSA, F.O., ALQUINI, Y., YATES, M.G., RIGO, L.U., STEFFENS, M.B., CHUBATSU, L.S., SOUZA, E.M. Root colonization systemic spreading and contribution of *Herbaspirillum seropedicae* to growth of rice seedling. **Symbiosis**, v.35, p.261-270, 2003b.

RONNING, C.M.; STEGALKINA, S.S.; ASCENZI, R.A.; BOUGRI, O., HART, A.L., UTTERBACH, T.R.; VANAKEN, S.E.; RIEDMULLER, S.B., WHITE, J.A., CHO, J.; et al. Comparative analyses of potato expressed sequence tag libraries. **Plant Physiol** v.131, p. 419–429, 2003.

ROUSE, D.; MACKAY, P.; STIMBERG, P.; ESTELLE, M.; LEYSER, O. Changes in auxin response from mutations in an *Aux/IAA* gene. **Science**. v. 279, p. 1371-1373, 1998.

SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. **Molecular cloning: a laboratory manual**. 2ed. New York, Cold Spring Harbor Laboratory Press, 1989.

SANO, Y.; FUJII, T.; IYAMA, S.; HIROTA, Y.; KOMAGATA, K. Nitrogen fixation in rizosphere of cultivated and wild rice strains. **Crop Sci.** v.21, p. 758-761, 1981

SASAKI, T.; BURR, B. International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. **Curr. Opin. Plant Biol.** v.3, p.138—141, 2000.

SCHULTZE, M.; A. KONDOROSI. Regulation of symbiotic root nodule development. **Annu. Rev. Genet.** v.32, p.33-57. 1998.

SHAW, L.J.; MORRIS, P.; HOOKER, J.E. Perception and modification of plant flavonoid signals by rhizosphere microorganisms. **Env. Microbiology.** V. 8, p. 1867-1880, 2006.

SHIMONO, M.; YAHZAKI, J.; NAKAMURA, K.; KISHIMOTO, N.; KIKUCHI, S.; IWANO, M.; YAMAMOTO, K.; SAKATA, K.; SASAKI, T.; NISHIGUCHI, M. cDNA microarray analysis of gene expression in rice plants treated with probenazole, a chemical inducer of disease resistance. **J. Gen. Pathol.** v. 69, p. 76-82, 2003.

SHRESTHA, R. K., E J. K. LADHA. Genotypic variation in promotion of rice nitrogen fixation as determined by ¹⁵N dilution. **Soil Sci. Am. J.** v.60, p.1815-1821, 1996.

SIMPSON, F. B.; BURRIS, R. H. A nitrogen pressure of 50 atmospheres does not prevent evolution of hidrogen by nitrogenase. **Science**, v. 224, p. 1095-1096, 1984.

SOUZA, E. M.; PEDROSA, F. O.; MACHADO, H. B.; DRUMMOND, M.; YATES, M. G. The N-terminus of the NifA protein of *Herbaspirillum seropedicae* is probably involved in sensing of ammonia. **In I. A. Tikhonovich, N. A. Provorov, V. I. Romanov, and W. E. Newton (ed.), Nitrogen fixation: fundamentals and applications. Kluwer Academic Publishers, Dordrecht, The Netherlands. p. 260. 1995.**

SPAINK, H. P. Root nodulation and infection factors produced by rhizobial bacteria. **Annu. Rev. Microbiol.** v.54, p.257-288, 2000.

SPRENT, J.I.; JAMES, E.K. N₂ fixation by endophytic bacteria: question of entry and operation. **In Azospirillum VI and Related Microorganisms**. Eds. Fendrik, I.; del Gallo, M.; VANDERLEYDEN, J.; de ZAMAROCZY, M. p. 15-30. Springer, Berlin, Heidelberg, Germany, 1995.

STERCK, L.; ROMBAUTS, S.; JANSSON, S. et al. EST data suggest that poplar is an ancient polyploid. **New Phytologist**. v.167, p.165-170, 2005.

STOLTZFUS, J.R.; SO, R.; MALARVITHI, P.P.; LADHA, J.K.; de BRUIJN, F.J. Isolation of endophytic bacteria from rice and assessment of their potential for supplying rice with biologically fixed nitrogen. **Plant and Soil**. v.194, p.25-36, 1997.

TIAN, Q.; REED, J. Control of auxin-regulated root development by *Arabidopsis thaliana* SHY/IAA3 gene. **Development**. v. 126, p.711-721, 1999.

TIWARI, S.B.; HAGEN, G.; GUILFOYLE, T. The roles of auxin response factor domains in auxin-responsive transcription. **The Plant Cell**. v. 15, p 533-543, 2003.

TIWARI, S.B.; WANG, X.; HAGEN, G.; GUILFOYLE, T.J. AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. **The Plant Cell**. v. 13, p. 2809-2822, 2001.

UEDA, T., Y. SUGA, N. YAHIRO, E T. MATSUGUCHI. Remarkable N₂-fixing bacterial diversity detected in rice roots by molecular evolutionary analysis of *nifH* gene sequences. **J. Bacteriol.** v.177, p.1414-1417, 1995.

ULMASOV, T.; HAGEN, G.; GUILFOYLE, T. J. ARF1, a transcription factor that binds to auxin response elements. **Science**. v. 276, p. 1865-1868, 1999.

URQUIAGA, S., K. H. S. CRUZ, E R. M. BODDEY. Contribution of nitrogen fixation to sugar cane: nitrogen-15 and nitrogen balance estimates. **Soil Sci. Soc. Am. J.** v. 56, p.105-114, 1992.

VAN BUREN, A.M.; ANDRE, C.; ISHIMARU, C.A. Biological control of the bacterial ring root pathogen by endophytic bacteria isolated from potatoe. **Phytopathology**. v.83, p.1406, 1993.

VAN PEER, R.; SCHIPPERS, B. Plant growth response to bacterization with selected *Pseudomonas*: effect of soil organic matter content, number of rhizosphere bacteria and timing of inoculation. **Soil Biol. Biochem.** v.20, p.45-49, 1989.

VENTELON-DEBOUT, M.; NGUYEN, T.T.H.; WISSOCQ, A.; et al. Analysis of the transcripcional response to Rice Yellow Mottle Virus infection in *Oryza sativa indica*

and *japonica* cultivars. **Molecular Genetics and Genomics**, v.270, p.253-262, 2003.

VIEIRA, J.; MESSING, J. The pUC plasmids, an M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primer. **Gene**, v.19, p. 259-268, 1982.

WANG, D.; PEI, K.; FU, Y.; SUN, Z.; LI, S.; LIU, H.; TANG, K.; HAN, B.; TAO, Y. Genome-wide analysis of the *auxin response factors* (ARF) gene family in rice (*Oryza sativa*). v. 394, p. 13-24, 2007.

WASDE (World Agricultural Supply and Demand Estimates):
<http://usda.mannlib.cornell.edu/reports/waobr/wasde-bb/2001/wasde377.txt>

WATSON, L. AND DALLWITZ, M.J. (1992 onwards). The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, and references. <http://delta-intkey.com>

WONG, H.L.; SAKAMOTO, T.; KAWASAKI, T.; UMEMURA, K.; SHIMAMOTO, K. Down-regulation of metallothionein, a reactive oxygen scavenger, by the small GTPase OsRac1 in rice. **Plant Physiol.** v.135, p.1447-1456, 2004.

WU, J.; MAEHARA, T.; SHIMOKAWA, T. et al. A comprehensive rice transcript map containing 6591 expressed sequence tag sites. **The Plant Cell**. v.14, p.525-535, 2002.

WU, P., G. ZHANG, J. K. LADHA, S. R. MCCOUCH, E. N. HUANG. Molecular-marker-facilitated investigation on the ability to stimulate N₂ fixation in the rhizosphere by irrigated rice plants. **Theor. Appl. Genet.** v.91, p.1171-1183, 1995.

YU, J., WANG, J., LIU, W., LI, S., LI, H., et al. The genomes of *Oryza sativa*: A history of duplication. **Plos Biol.** v.3, n.2, p.266-281, 2005.

YU, L.; UMEDA, M.; LIU, J.; ZHAO, N.; UCHIMIYA, H. A novel MT gene of rice plants is strongly expressed in the node portion of the stem. **Gene**. v. 206, p.29-35, 1998.

ZEGZOUTI, H.; JONES, B.; FRASSE, P.; MARTY, C.; MAITRE, B.; LATCHÉ, A.; PECH, J.; BOUZAYEN, M. Ethylene-regulated gene expression in tomato fruit: characterization of novel ethylene-responsive and ripening-related genes isolated by differential display. **The Plant Journal**. v.18, p.589-600, 1999.

ZHANG, D.; CHOI, D.W.; WANAMAKER, S.; et al. Construction and evaluation of cDNA libraries for large-scale expressed sequence tag sequencing in wheat (*Triticum aestivum* L.). **Genetics**. v. 168, p.595-608, 2004.

ZHOU, G.; XU, Y.; LI, J.; YANG, L.; LIU, J. Molecular Analyses of the metallothionein gene family in rice (*Oryza sativa* L.). **J. Biochem. Mol.** v.39, p.595-606, 2006.

ZHU, G; ZHUANG, C.; WANG, Y; JIANG, L.; PENG, X. Differential expresión of rice genes under different nitrogen forms and their relationship with sulfur metabolismo. **Journal of Integrative Plant Biology**. v. 48, p.1177-1184, 2006.